

DRAFT 514

012

CLASS

Docket No.: 1101-209-999

Serial No.: 09/079,819

Inventor(s): ALVAREZ et al.

Title: "RANDOM PEPTIDES THAT BIND TO GASTROINTESTINAL TRACT (GIT) TRANSPORT RECEPTORS AND RELATED METHODS"

20	40	60
MGMSKSHSFFGYPLSIFFI	VNEFCERFSYYGMRAIL	LYFTNFISWDDNLSTAIYHTFV
80	100	120
ALCYLTPILGALIADSWLGK	FKTIVSLSIVYTIGQAVTSV	SSINDLTDHNDGTPDPLPV
140	160	180
HVLSLIGLALIALGTGGIK	PCVSAFGGDQFEEGQEKQRN	RFFSIFYLAINAGSLLSTII
200	220	240
TPMLRVQQCGIHSKQACYPL	AFGVPAALMAVALIVFVLGS	GMYYKKFKPQGNIMGKVAKCI
260	280	300
GFAIKNRFRHRSKAFPKREH	WLDWAKEKYDERLISQIKMV	TRVMFLYIPLPMFWALFDQQ
320	340	360
GSRWTLQATTMSGKIGALEI	QPDQMOTVNAILIVIMVPIF	DAVLYPLIAKCGFNFTSLKK
380	400	420
MAVGMVLASMAFVAAIVQV	EIDKTLPVFPKGNEVQIKVL	NIGNNTMNISLPGEMVTLGP
440	460	480
MSQTNAFMFTDVNKLTRINI	SSPGSPVTAVTDDFKQGQRH	TLLVWAPNHYQVVKDGLNQK
500	520	540
PEKGENGIRFVNTFNELITI	TMSGKVYANISSYNASTYQF	FPSGIKGFTISSTEIPPQCQ
560	580	600
PNFNTFYLEFGSAYTYIVQR	KNDSCPEVKVFEDISANTVN	MALQIPQYFLLTCGEVVFVS
620	640	660
TGLEFSYSQAPSNMKSVLQA	GWLLTVAVGNIIVLIVAGAG	QFSKQWAEYILFAALLLVVC
680	700	708
VIFAIMARFYTYINPAEIEA	QFDEDEKKNRLEKSNPYFMS	GANSQKQM

FIG.1

DRAFT

1	gaattccgctc	tcgaccactg	aatggaagaa	aaggactttt	aaccaccatt	ttgtgactta
61	cagaaaggaa	tttgaataaa	gaaaactatg	atacttcagg	cccatcttca	ctccctgtgt
			M	I L Q	A H L	H S L C
121	cttcttatgc	tttatttggc	aactggatat	ggccaagagg	ggaagttag	tggacccctg
	L L M	L Y L	A T G Y	G Q E	G K F	S G P L
181	aaacccatga	cattttctat	ttatgaaggc	caagaaccga	gtcaaattat	attccagttt
	K P M	T F S	I Y E G	Q E P	S Q I	I F Q F
241	aaggccaatc	ctcctgctgt	gacttttgaa	ctaactgggg	agacagacaa	catatttgtg
	K A N	P P A	V T F E	L T G	E T D	N I F V
301	atagaacggg	agggacttct	gtattacaac	agagccttgg	acagggaaac	aagatctact
	I E R	E G L	L Y Y N	R A L	D R E	T R S T
361	cacaatctcc	aggttgcagc	cctggacgct	aatggaatta	tagtggaggg	tccagtcctt
	H N L	Q V A	A L D A	N G I	I V E	G P V P
421	atcaccatag	aagtgaagga	catcaacgac	aatcgaccca	cgtttctcca	gtcaaagtac
	I T I	E V K	D I N D	N R P	T F L	Q S K Y
481	gaaggctcag	taaggcagaa	ctctcgccca	ggaaagccct	tcttgatatg	caatgccaca
	E G S	V R Q	N S R P	G K P	F L Y	V N A T
541	gacctggatg	atccggccac	tcccaatggc	cagctttatt	accagattgt	catccagctt
	D L D	D P A	T P N G	Q L Y	Y Q I	V I Q L
601	cccatgatca	acaatgtcat	gtactttcag	atcaacaaca	aaacgggagc	catctctctt
	P M I	N N V	M Y F Q	I N N	K T G	A I S L
661	acccgagagg	gatctcagga	attgaatcct	gctaagaatc	cttcctataa	tctggtgatc
	T R E	G S Q	E L N P	A K N	P S Y	N L V I
721	tcagtgaagg	acatgggagg	ccagagtgag	aattccttca	gtgataccac	atctgtggat
	S V K	D M G	G Q S E	N S F	S D T	T S V D
781	atcatagtga	cagagaatat	ttggaaagca	ccaaaacctg	tggagatggg	ggaaaactca
	I I V	T E N	I W K A	P K P	V E M	V E N S
841	actgatcctc	accccatcaa	aatcactcag	gtgcggtgga	atgatcccgg	tgcacaatat
	T D P	H P I	K I T Q	V R W	N D P	G A Q Y
901	tccttagttg	acaaagagaa	gctgccaaaga	ttcccatttt	caattgacca	ggaaggagat
	S L V	D K E	K L P R	F P F	S I D	Q E G D
961	atttacgtga	ctcagccctt	ggaccgagaa	gaaaaggatg	catatgtttt	ttatgcagtt
	I Y V	T Q P	L D R E	E K D	A Y V	F Y A V
1021	gcaaaggatg	agtacggaaa	accactttca	tatccgctgg	aaattcatgt	aaaagttaaa
	A K D	E Y G	K P L S	Y P L	E I H	V K V K
1081	gatattaatg	ataatccacc	tacatgtccg	tcaccagtaa	ccgtatttga	gggccaggag
	D I N	D N P	P T C P	S P V	T V F	E V Q E
1141	aatgaacgac	tgggtaacag	tatcgggacc	cttactgcac	atgacaggga	tgaagaaaat
	N E R	L G N	S I G T	L T A	H D R	D E E N
1201	actgccaaaca	gttttctaaa	ctacaggatt	gtggagcaaa	ctcccaaact	tcccatggat
	T A N	S F L	N Y R I	V E Q	T P K	L P M D

FIG. 2A

1261 ggactcttcc taatccaaac ctatgctgga atgttacagt tagctaaaca gtccttgaag
G L F L I Q T Y A G M L Q L A K Q S L K
1321 aagcaagata ctctcagta caacttaacg atagaggtgt ctgacaaaga tttcaagacc
K Q D T P Q Y N L T I E V S D K D F K T
1381 ctttgttttg tgcaaatcaa cgttattgat atcaatgata agatcccat ctttgaaaaa
L C F V Q I N V I D I N D Q I P I F E K
1441 tcagattatg gaaacctgac tcttgctgaa gacacaaaca ttgggtccac catcttaacc
S D Y G N L T L A E D T N I G S T I L T
1501 atccaggcca ctgatgctga tgagccattt actgggagtt ctaaaattct gtatcatatc
I Q A T D A D E P F T G S S K I L Y H I
1561 ataaagggag acagtgaggg acgcctgggg gttgacacag atccccatac caacaccgga
I K G D S E G R L G V D T D P H T N T G
1621 tatgtcataa ttaaaaagcc tcttgatttt gaaacagcag ctgtttccaa cattgtgttc
Y V I I K K P L D F E T A A V S N I V F
1681 aaagcagaaa atcctgagcc tctagtgttt ggtgtgaagt acaatgcaag ttcttttggc
K A E N P E P L V F G V K Y N A S S F A
1741 aagttcacgc ttattgtgac agatgtgaat gaagcacctc aattttccca acacgtattc
K F T L I V T D V N E A P Q F S Q H V F
1801 caagcgaag tcagtgagga tgtagctata ggcactaaag tgggcaatgt gactgccaag
Q A K V S E D V A I G T K V G N V T A K
1861 gatccagaag gtctggacat aagctattca ctgaggggag acacaagagg ttggcttaaa
D P E G L D I S Y S L R G D T R G W L K
1921 attgaccacg tgactggtga gatctttagt gtggctccat tggacagaga agccggaagt
I D H V T G E I F S V A P L D R E A G S
1981 ccatatcggg tacaagtggg ggccacagaa gtaggggggt cttccttaag ctctgtgtca
P Y R V Q V V A T E V G G S S L S S V S
2041 gatttccacc tgatccttat ggatgtgaat gacaaccctc ccaggctagc caaggactac
E F H L I L M D V N D N P P R L A K D Y
2101 acgggcttgt tcttctgcca tcccctcagt gcacctggaa gtctcatttt cgaggctact
T G L F F C H P L S A P G S L I F E A T
2161 gatgatgata agcacttatt tcggggctcc cattttacat tttccctcgg cagtggaagc
D D D Q H L F R G P H F T F S L G S G S
2221 ttacaaaacg actgggaagt ttccaaaatc aatgggtactc atgcccgact gtctaccagg
L Q N D W E V S K I N G T H A R L S T R
2281 cacacagact ttgaggagag ggcgtatgtc gtcttgatcc gcatcaatga tgggggtcgg
H T D F E E R A Y V V L I R I N D G G R
2341 ccacccttgg aaggcattgt ttctttacca gttacattct gcagttgtgt ggaaggaagt
P P L E G I V S L P V T F C S C V E G S
2401 tgtttccggc cagcaggtca ccagactggg ataccactg tgggcatggc agttggtata
C F R P A G H Q T G I P T V G M A V G I

FIG.2B

2461 ctgctgacca cccttctggt gattggtata attttagcag ttgtgtttat ccgcataaag
L L T T L L V I G I I L A V V F I R I K
2521 aaggataaag gcaaagataa tggtgaaagt gctcaagcat ctgaagtcaa acctctgaga
K D K G K D N V E S A Q A S E V K P L R
2581 agctgaattt gaaaaggaat gtttgeattt atatagcaag tgctatttca gcaacaacca
S

2641 tctcatccta ttacttttca tctaacgtgc attataattt tttaaacaga tattccctct
2701 tgtcctttta tatttgctaa atatttcttt tttagagggtg agtcttgctc tgtcgcccag
2761 gctggagtac agtggtgtga tcccagctca ctgcaacctc cgcctcctgg gttcacatga
2821 ttctcctgcc tcagcttcct aagtagctgg gtttacaggc acccaccacc atgcccagct
2881 aatttttgta tttttaatag agacggggtt tgcgccatttg gccaggatgg tcttgeactc
2941 ctgacgtcaa gtgatctgcc tgccttggtc tcccaatata ggcatgaacc actgcacca
3001 cctacttaga tatttcatgt gctatagaca ttagagagat ttttcatttt tccatgacat
3061 ttttcctctc tgcaaatggc ttagctactt gtgtttttcc cttttggggc aagacagact
3121 cattaaatat tctgtacatt ttttctttat caaggagata tatcagtgtt gtctcataga
3181 actgcctgga ttccatttat gttttttctg attccatcct gtgtcccctt catccttgac
3241 tcctttggta tttcactgaa tttcaaacad ttgtcagaga agaaaaaagt gaggactcag
3301 gaaaaataaa taaataaaaag aacagccttt tgcggccgcg aattc

FIG.2C

773
CLASS
DRAFT

Docket No.: 1101-209-999

Serial No.: 09/079,819

Inventor(s): ALVAREZ et al.

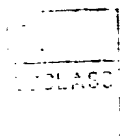
Title: "RANDOM PEPTIDES THAT BIND TO GASTROINTESTINAL
TRACT (GIT) TRANSPORT RECEPTORS AND RELATED METHODS"

20	40	60
MARKKFSGLEISLIVLFVIV	TIIAIALIVVLATKTPAVDE	ISDSTSTPATTRVTTNPSDS
80	100	120
GKCPNVLNDPVNVRINCIPE	QFPTEGICAQRGCCWRPWND	SLIPWCFFVDNHGYNVQDMT
140	160	180
TTSIGVEAKLNRI PSPTLFG	NDINSVLFTTQNQTPNRF	KITDPNNRRYEVPHQYVKEF
200	220	240
TGPTVSDTL YDKVAQNPF	IQVIRKSNGKTLFDTSIGPL	VYSDQYLQISARLP SDYIYG
260	280	300
IGE QVHKRFRHDL SWKTWPI	FTRDQLPGDNNNNLYGHQTF	FMCIEDTSGKSFGVFLMNSN
320	340	360
AMEIFIQPTPIV TYRVTGGI	LD FYILLGDTPEQVVQQYQQ	LVGLPAMPAYWNLGFQLSRW
380	400	420
NYKSLDVVKEVVRNREAGI	PFDTQVTDIDYMEDKKDFTY	DQVAFNGLPQFVQDLHDHGQ
440	460	480
KYVIILDPAISIGRRANGTT	YATYERGNTQHVWINESDGS	TPIIGEVWPGLTVYPDF TNP
500	520	540
NCIDWWANEC SIFHQEVQYD	GLWIDMNEVSSFIQGSTKGC	NVNKLNYPPFTPDILDKLMY
560	580	600
SKTICMDAVQNWGKQYDVHS	LYGYSMAIATEQAVQKVFPN	KRSFILTRSTFAGSGRHA AH
620	640	660
WLGDN TASWEQMEWSITGML	EFSLFGIPLVGADICGFVAE	TTEELCRRWMQLGAFYPFSR
680	700	720
NHNSDGYEHQDPAFFGQNSL	LVKSSRQYL TIRYTL LPFLY	TLFYKAHVFGETVARPVLHE
740	760	780
FYEDTNSWIEDTEFLWGPAL	LITPVLKQGADTVSAYIPDA	IWYDYESGAKRPWRKQRVDM
800	820	840
YLPADKIGLHLRGGYIPIQ	EPDVTTTASRKNPLGLIVAL	GENNTAKGDFFWDDGETKDT
860	880	900
IQNGNYILYTF SVSNNTLDI	VCTHSSYQEGTTLAFQTVKI	LGLTDSVTEVRVAENNQPMN
920	940	960
AHSNFTYDASNQVLLIADLK	LNLGRNFSVQWNQIFSENER	FNCYPDADLATEQKCTQRGC
980	1000	1020
VWRTGSSLSKAPECYFPRQD	NSYSVNSARYSSMGITADLQ	LNTANARIKLPSDPISTLRV
1040	1060	1080
EVKYHKNDMLQFKIYDPQKK	RYEVPVPLNIPTTPISTYED	RLYDVEIKENPFGIQIRRRS
1100	1120	1140
SGRVIWDSWLP GFAFNDQFI	QISTRLPSEYIYGFGVEHT	AFKRDLNWNTWGMFTRDQPP
1160	1180	1200
GYKLNSYGFHPYYMALEEEG	NAHGVFLLNSNAMDVTFQPT	PALTYRTVGGILDYFMLG P
1220	1240	1260
TPQVATKQYHEVIGHPVMPA	YWALGFQLCRYGYANTSEVR	ELYDAMVAANIPYDVQYTDI

FIG.3A

1280	1300	1320
DYMERQLDFTIGEAFQDLPO	FVDKIRGEGMRYIIILDPAI	SGNETKTYPAFERGQQNDVF
1340	1360	1380
VKWPNTNDICWAKVWPDLPN	ITIDKTLTEDEAVNASRAHV	AFPDFFRTSTAEEWAREIVD
1400	1420	1440
FYNEKMKFDGLWIDMNEPSS	FVNGTTTNQCRNDELNYPPY	FPELTKRTDGLHFRTICMEA
1460	1480	1500
EQILSDGTSVLHYDVHNLVG	WSQMKPTHDALQKTTGKRG	VISRSTYPTSGRWGGHWLGD
1520	1540	1560
NYARWDNMDKSIIGMMEFSL	FGISYTGADICGFFNNSEYH	LCTRWMQLGAFYPYSRNHNI
1580	1600	1620
ANTRRQDPASWNETFAEMSR	NILNIRYTLTPYFYTMHEI	HANGGTVIRPLLHEFFDEKP
1640	1660	1680
TWDIFKQFLWGPAFMVTPVL	EPYVQTVNAYVPNARWFDYH	TGKDIGVRGQFQTFNASYDT
1700	1720	1740
INLHVRGGHILPCQEPAQNT	FYSRQKHKMLIVAADDNQMA	QGSLFWDDGESIDTYERDLY
1760	1780	1800
LSVQFNLNQTTLTSTILKRG	YINKSETRLGSLHVWGKGT	PVNAVTLTYNGKNKNSLPFNE
1820	1827	
DTTNMILRIDLTTHNVTLEE	PIEINWS	

FIG.3B



```

1 gccttactgc aggaaggcac tccgaagaca taagtcggtg agacatggct gaagataaaa
                                M A E D K
61 gcaagagaga ctccatcgag atgagtatga agggatgcca gacaaacaac gggtttgtcc
  S K R D S I E M S M K G C Q T N N G F V
121 ataatgaaga cattctggag cagaccccgg atccaggcag ctcaacagac aacctgaagc
  H N E D I L E Q T P D P G S S T D N L K
181 acagcaccag gggcatcctt ggctcccagg agcccgactt caagggcgctc cagccctatg
  H S T R G I L G S Q E P D F K G V Q P Y
241 cgggggatgcc caaggagggtg ctgttccagt tctctggcca ggcccgtac cgcatacctc
  A G M P K E V L F Q F S G Q A R Y R I P
301 gggagatcct cttctggctc acagtggctt ctgtgctggt gctcatcgcg gccaccatag
  R E I L F W L T V A S V L V L I A A T I
361 ccatcattgc cctctctcca aagtgcctag actggtggca ggaggggccc atgtaccaga
  A I I A L S P K C L D W W Q E G P M Y Q
421 tctaccaag gtctttcaag gacagtaaca aggatgggaa cggagatctg aaaggtattc
  I Y P R S F K D S N K D G N G D L K G I
481 aagataaact ggactacatc acagctttaa atataaaaac tgtttggatt acttcatttt
  Q D K L D Y I T A L N I K T V W I T S F
541 ataaatcgtc ccttaaagat ttcagatatg gtgttgaaga tttccgggaa gttgatccca
  Y K S S L K D F R Y G V E D F R E V D P
601 tttttggaac gatggaagat tttgagaatc tggttgcagc catacatgat aaaggtttaa
  I F G T M E D F E N L V A A I H D K G L
661 aattaatcat cgatttcata ccaaaccaca cgagtgataa acatatttgg tttcaattga
  K L I I D F I P N H T S D K H I W F Q L
721 gtcggacacg gacaggaaaa tatactgatt attatatctg gcatgactgt acccatgaaa
  S R T R T G K Y T D Y Y I W H D C T H E
781 atggcaaaac cattccaccc aacaactggt taagtgtgta tggaaactcc agttggcact
  N G K T I P P N N W L S V Y G N S S W H
841 ttgacgaagt gcgaaaccaa tgttattttc atcagtttat gaaagagcaa cctgatttaa
  F D E V R N Q C Y F H Q F M K E Q P D L
901 atttccgcaa tcctgatgtt caagaagaaa taaaagaaat tttacggttc tggctcacia
  N F R N P D V Q E E I K E I L R F W L T
961 aggggtgttga tggtttttagt ttggatgctg ttaaattcct cctagaagca aagcacctga
  K G V D G F S L D A V K F L L E A K H L

```

FIG.4A

1021 gagatgagat ccaagtaa at aagacccaaa tcccggacac ggtcacacaa tactcggagc
R D E I Q V N K T Q I P D T V T Q Y S E

1081 tgtaccatga cttcaccacc acgcaggtgg gaatgcacga cattgtccgc agcttccggc
L Y H D F T T T Q V G M H D I V R S F R

1141 agaccatgga ccaatacagc acggagcccc gcagatacag gttcatgggg actgaagcct
Q T M D Q Y S T E P G R Y R F M G T E A

1201 atgcagagag tattgacagg accgtgatgt actatggatt gccatttata caagaagctg
Y A E S I D R T V M Y Y G L P F I Q E A

1261 attttccatt caacaattac ctcagcatgc tagacactgt ttctgggaac agcgtgtatg
D F P F N N Y L S M L D T V S G N S V Y

1321 aggttatcac atcctggatg gaaaacatgc cagaaggaaa atggcctaac tggatgattg
E V I T S W M E N M P E G K W P N W M I

1381 gtggaccaga cagttcacgg ctgacttcgc gtttggggaa tcagtatgtc aacgtgatga
G G P D S S R L T S R L G N Q Y V N V M

1441 acatgcttct ttccacactc cctggaactc ctataactta ctatggagaa gaaattggaa
N M L L F T L P G T P I T Y Y G E E I G

1501 tgggaaatat tgtagccgca aatctcaatg aaagctatga tattastacc cttcgctcaa
M G N I V A A N L N E S Y D I N T L R S

1561 agtcaccaat gcagtgggac aatagttcaa atgctggtt ttctgaagct agtaacacct
K S P M Q W D N S S N A G F S E A S N T

1621 ggttacctac caattcagat taccacactg tgaatgttga tgtccaaaag actcagccca
W L P T N S D Y H T V N V D V Q K T Q P

1681 gatcggcttt gaagttatat caagatttaa gtctacttca tgccaatgag ctactcctca
R S A L K L Y Q D L S L L H A N E L L L

1741 acaggggctg gttttgccat ttgaggaatg acagccacta tgttgtgtac acaagagagc
N R G W F C H L R N D S H Y V V Y T R E

1801 tggatggcat cgacagaatc tttatcgtgg ttctgaattt tggagaatca acactgttaa
L D G I D R I F I V V L N F G E S T L L

1861 atctacataa tatgatttcg ggccttcccc ctaaaataag aataaggtta agtaccatt
N L H N M I S G L P A K I R I R L S T N

1921 ctgccgacaa aggcagtaaa gttgatacaa gtggcatttt tctggacaag ggagagggac
S A D K G S K V D T S G I F L D K G E G

1981 tcatttttga acacaacacg aagaatctcc ttcatcgcca aacagctttc agagatagat
L I F E H N T K N L L H R Q T A F R D R

2041 gctttgtttc caatcgagca tgctattcca gtgtactgaa catactgtat acctcgtgtt
C F V S N R A C Y S S V L N I L Y T S C

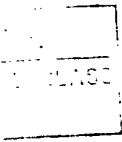
2101 aggcaccttt atgaagagat gaagacactg gcatttcagt gggattgtaa gcatttgtaa

2161 tagcttcatg tacagcatgc tgcttggtga acaatcatta attcttcgat atttctgtag

2221 cttgaatgta accgctttta gaaaggttct caaatgtttt gaaaaaata aaatgtttta

2281 aagt

FIG.4B



EXPRESSION OF PHAGE INSERTS AS GST FUSION

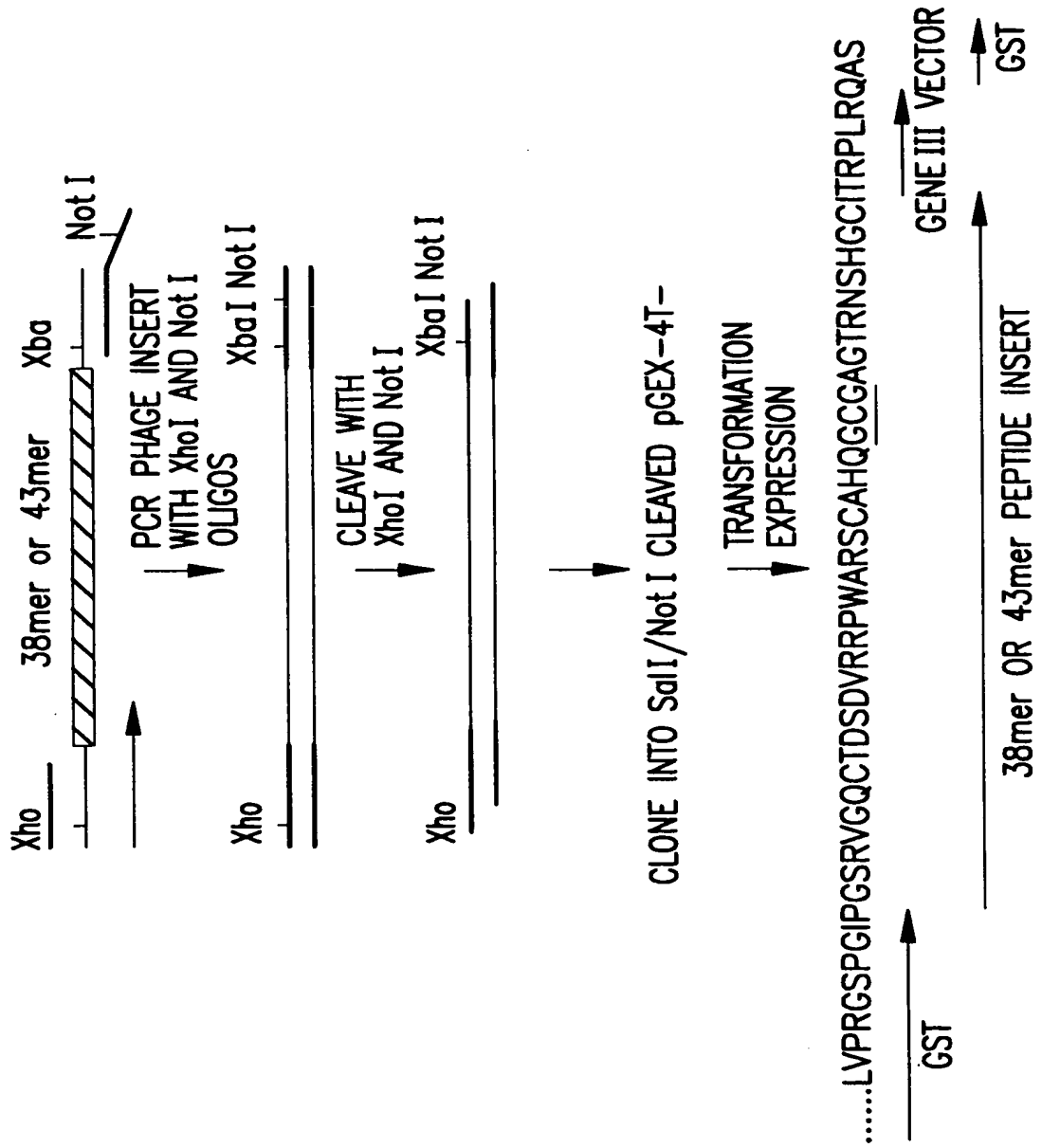


FIG. 5A

DRAFT

CLASS

P31	1	10	20	30	Clone #
	SARDSGPAEDGSRVRLNGVENANTRKSSRSNPRGRRHP				
	SARDSGPAEDGSRVRLNG				101
		DGSRAVRLNGVENANTRKSSR			102
			ENANTRKSSRSNPRGRRHP		103
			TRKSSRSNPRG		119
Pax2	1	10	20	30	Clone #
	STPPSREAYSRPYSDSDSDTNAKHSSHNRRLRTRSRPN				
	STPPSREAYSRPYSDSDSD				104
		SRPYSDSDSDTNAKHSSHNR			105
			TNAKHSSHNRRLRTRSRPN		106
DCX8	1	10	20	30	Clone #
	RYKHDIGCDAGVDKKSSSVRGCGAHSSPPRAGRGRGTMTMSRL				
	RYKHDIGCDAGVDKKSSSVRGCG				107
		GCDAGVDKKSSSVRGCGAHSSPPRA			108
			GAHSSPPRAGRGRGTMTMSRL		109

FIG.5B

P31	1	10	20	30	Clone #
	SARDSGPAEDGSRAVRLNGVENANTRKSSRSNPRGRRHP				
			ENANTRKSSRSNPRGRRHP		103
			ENANTRKSSR		110
			TRKSSRSNPRG		119
			RKSSRSNPRG		111
			SNPRGRRHP		112
Pax2	1	10	20	30	Clone #
	STPPSREAYSRPYSVDSDDTNAKHSSHNRRLRTRSRPN				
			TNAKHSSHNRRLRTRSRPN		106
			TNAKHSSH		113
			SSHNRRLRTR		114
			RRLRTRSRPN		115
SNi10	1	10	20	30	Clone#
	RVGQCTDSDVRRPWARSCAHQCGAGTRNSHGCITRPLRQASAH				
	RVGQCTDSDVRRPWARSCA				116
		VRRPWARSCAHQCGAGTRNS			117
			GTRNSHGCITRPLRQASAH		118

FIG.5C

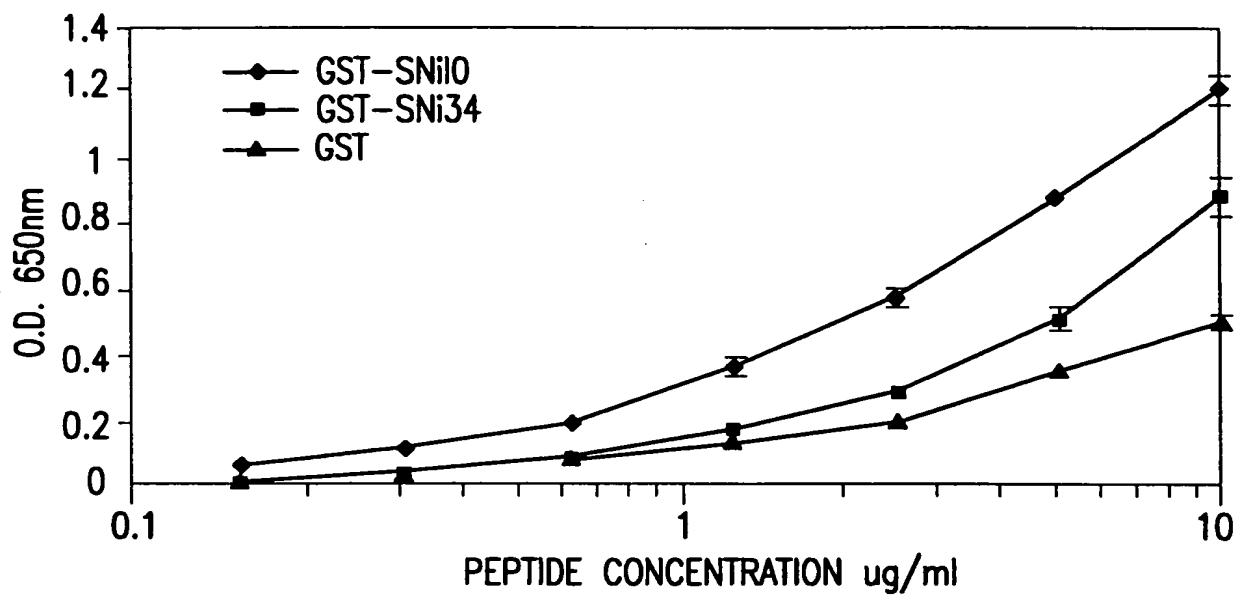


FIG.6A

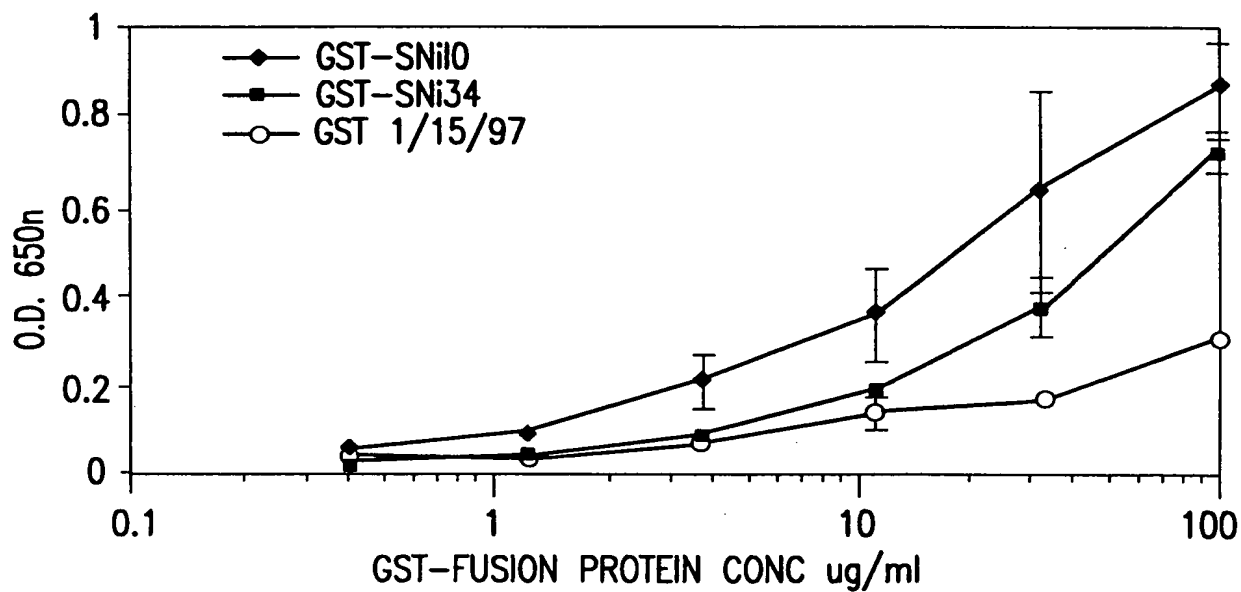


FIG.6B

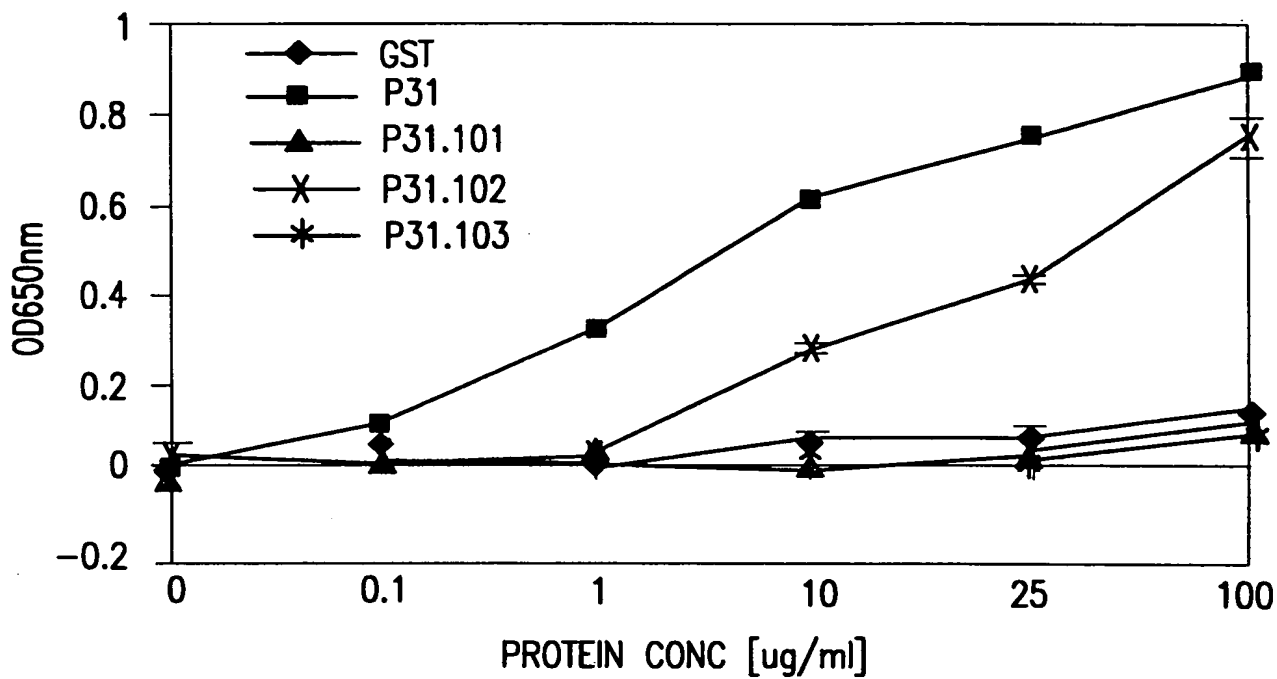


FIG. 7A

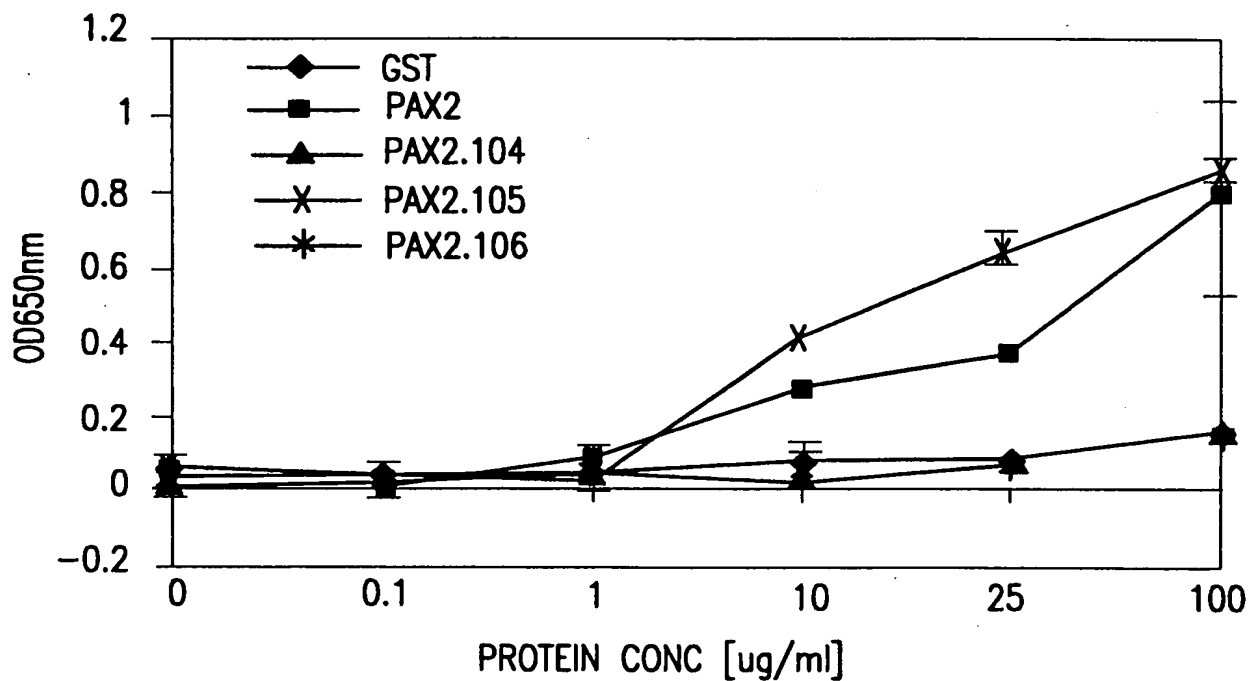


FIG. 7B

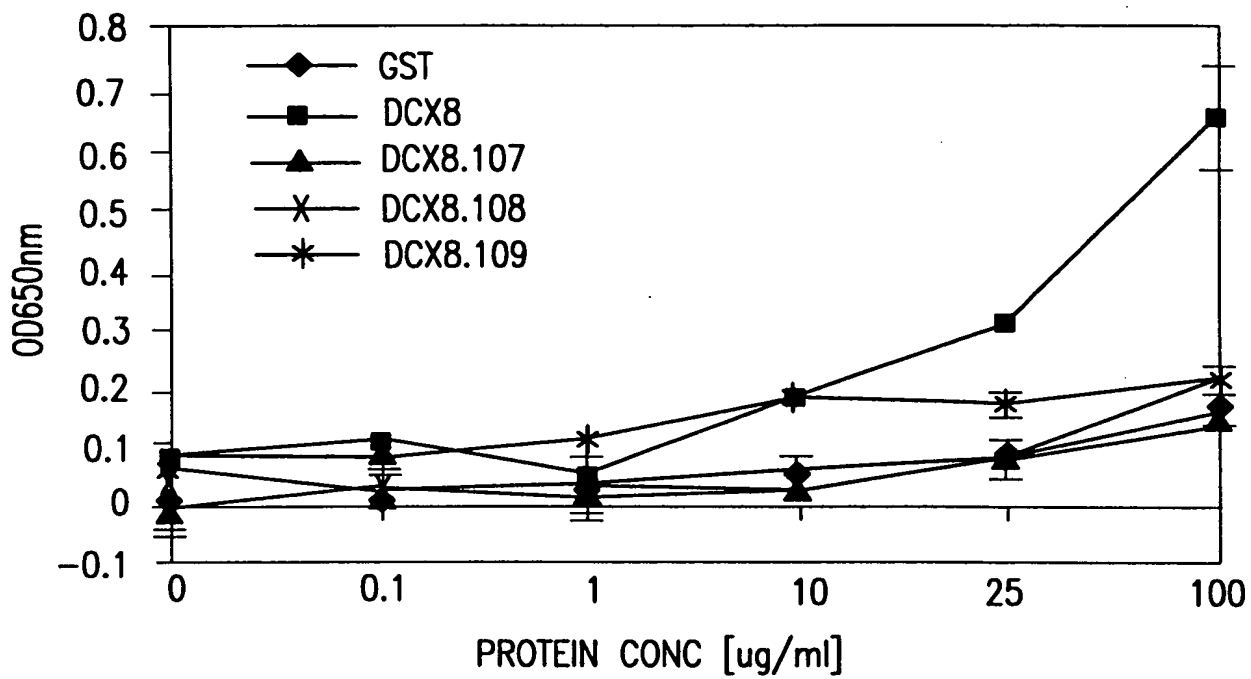


FIG. 7C

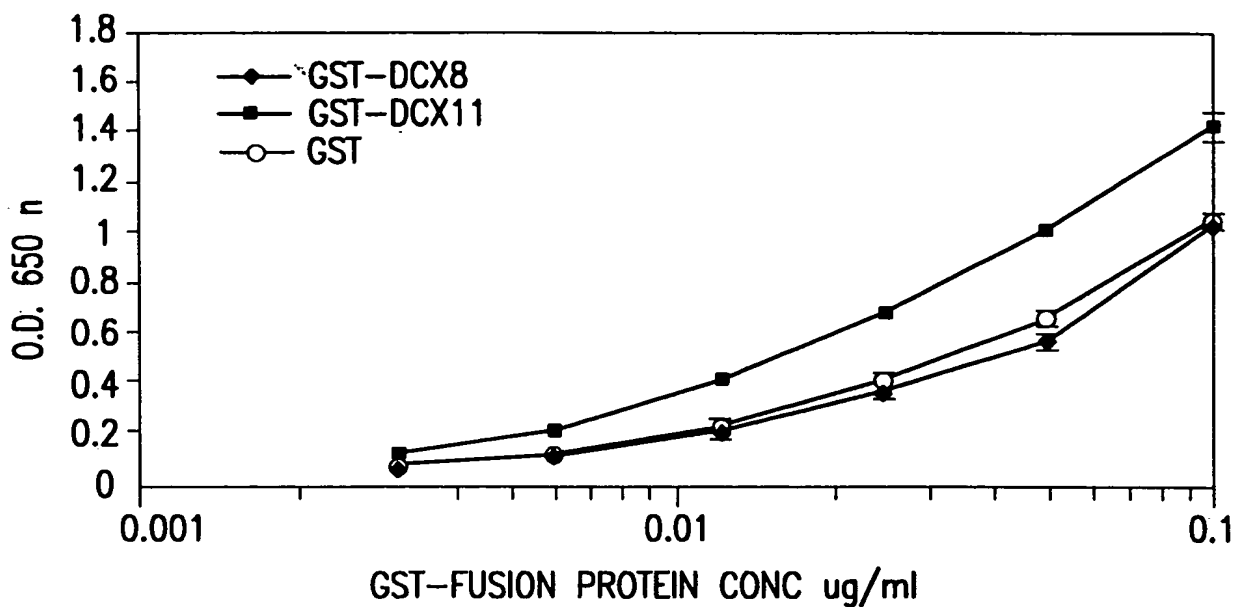
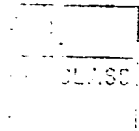


FIG.7D

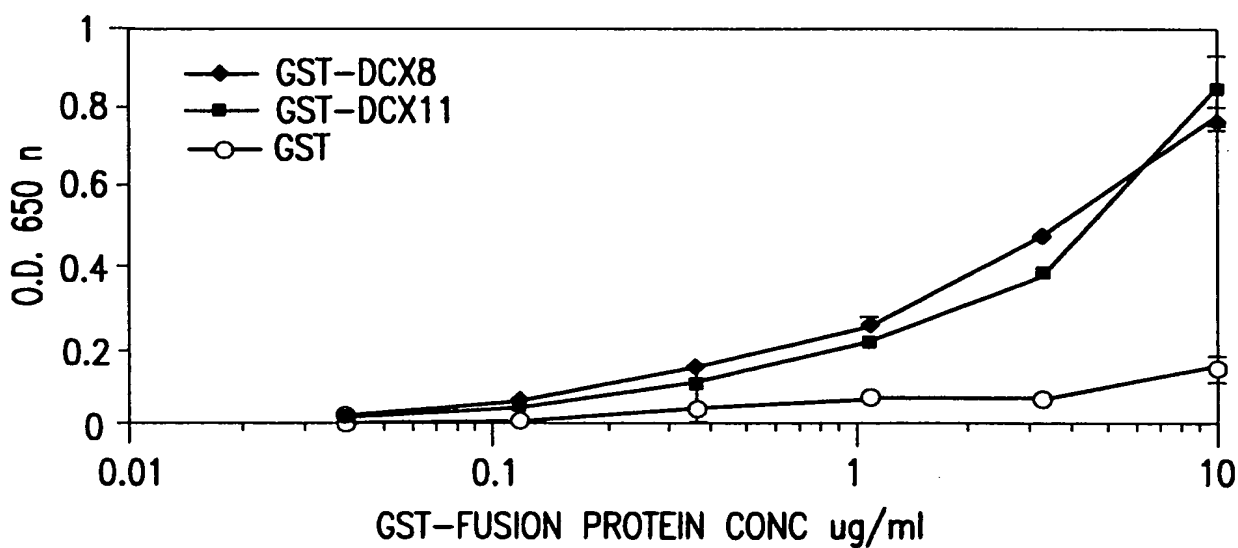


FIG.7E

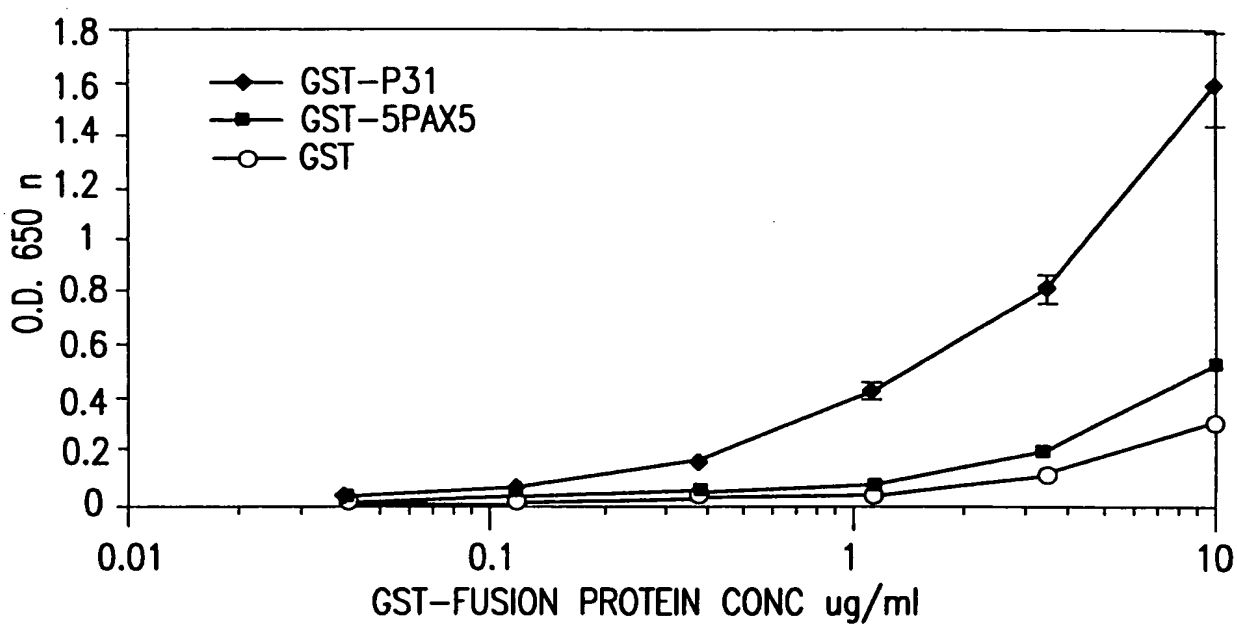


FIG.7F

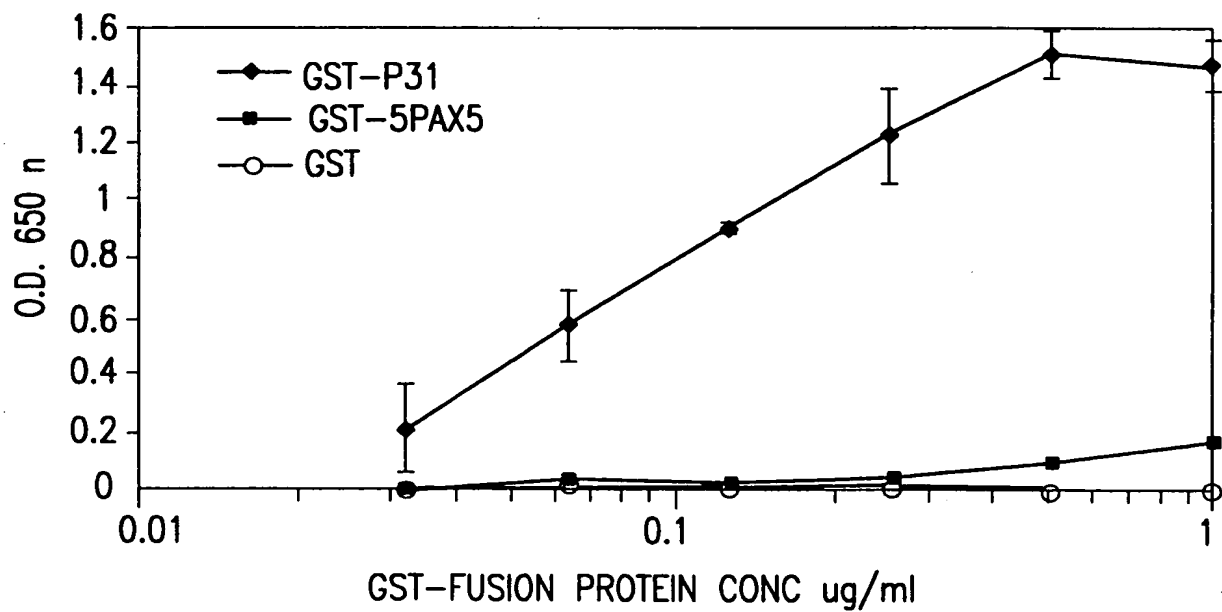


FIG.7G

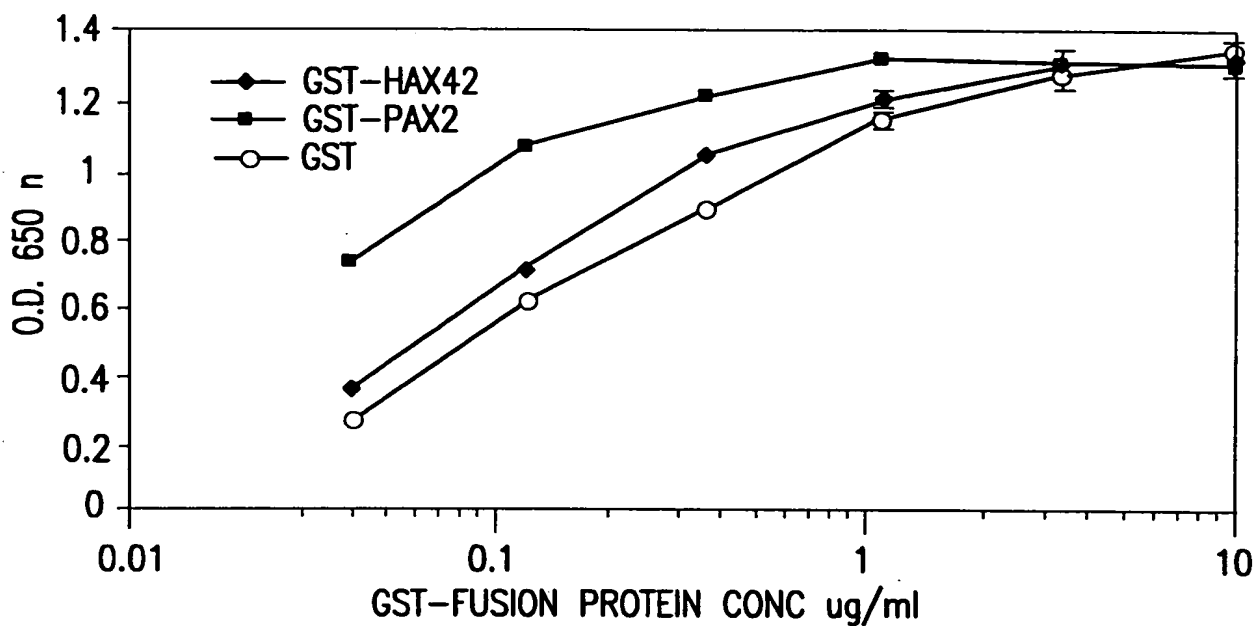
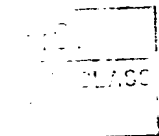


FIG. 7H

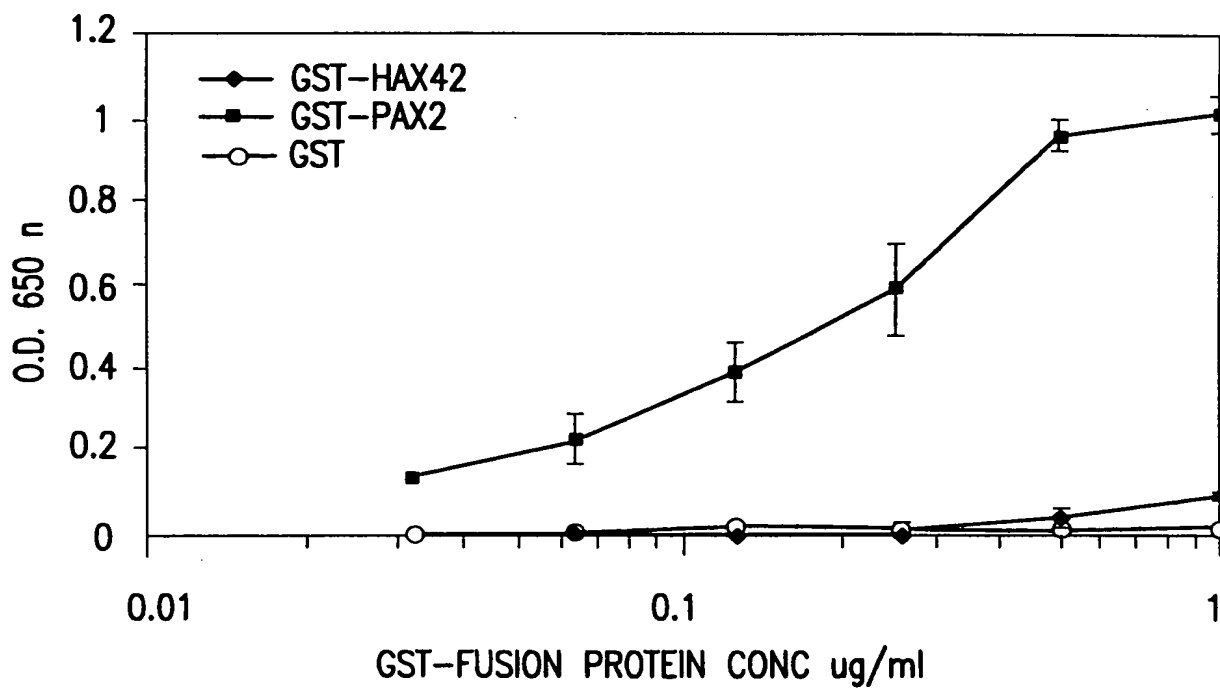


FIG. 7I

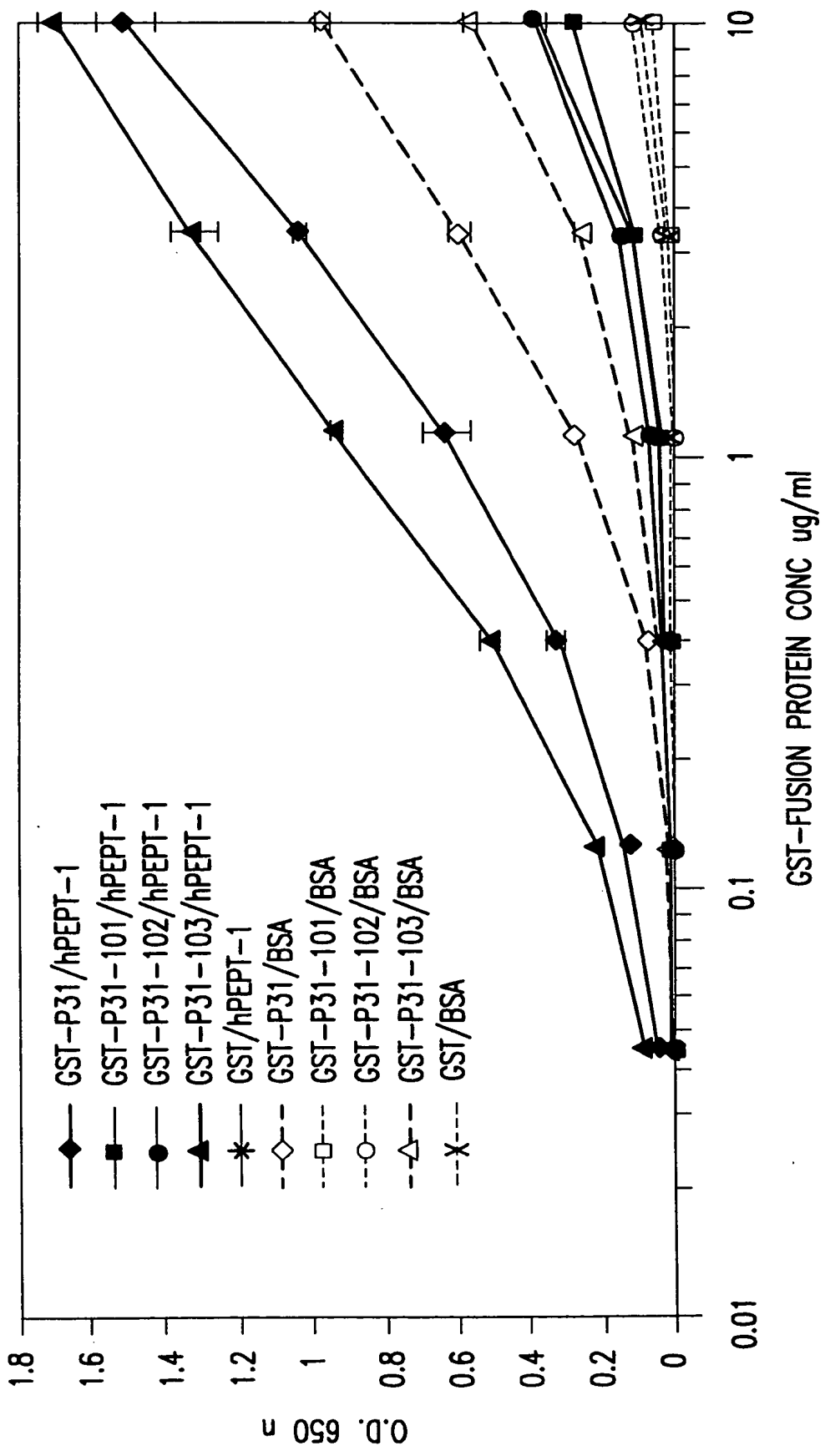


FIG. 7J

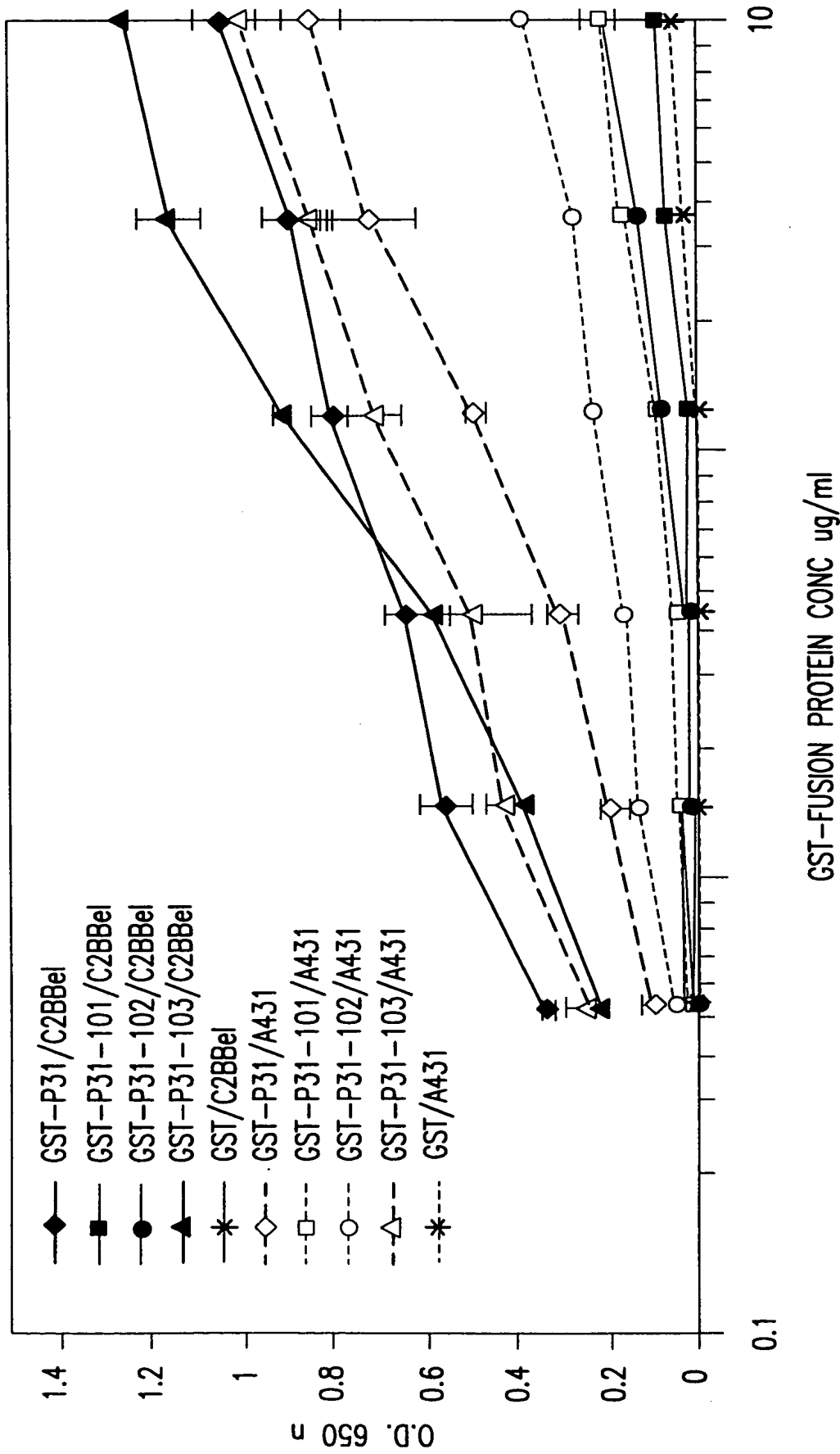


FIG. 7K

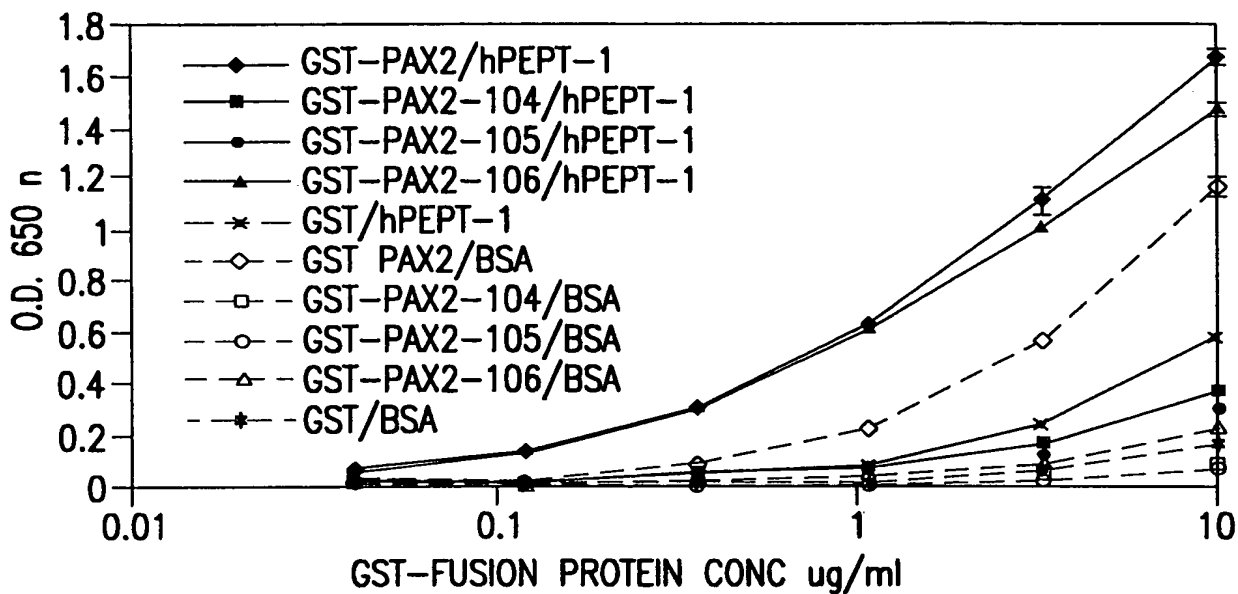


FIG. 7L

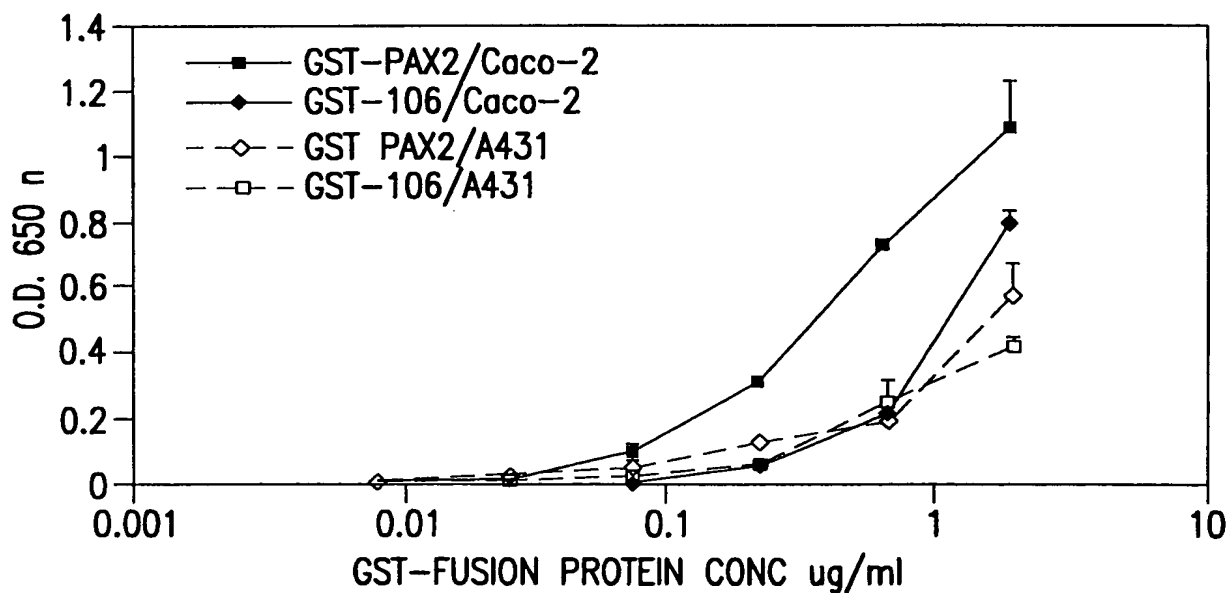


FIG. 7M

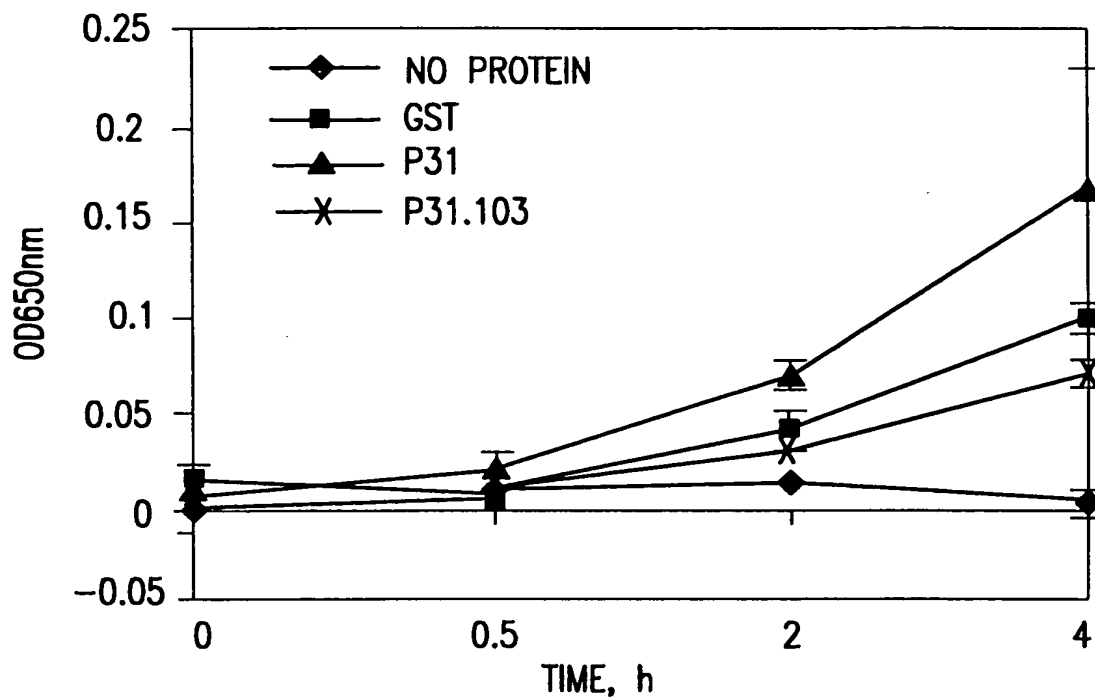


FIG.8A

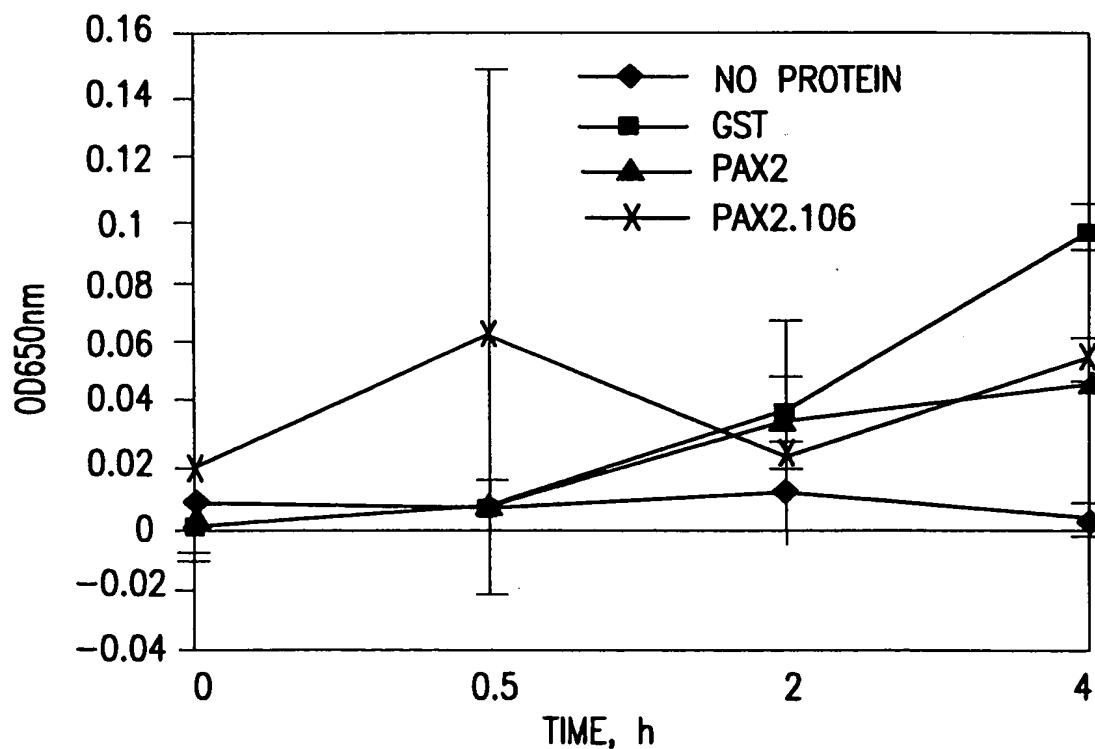


FIG.8B

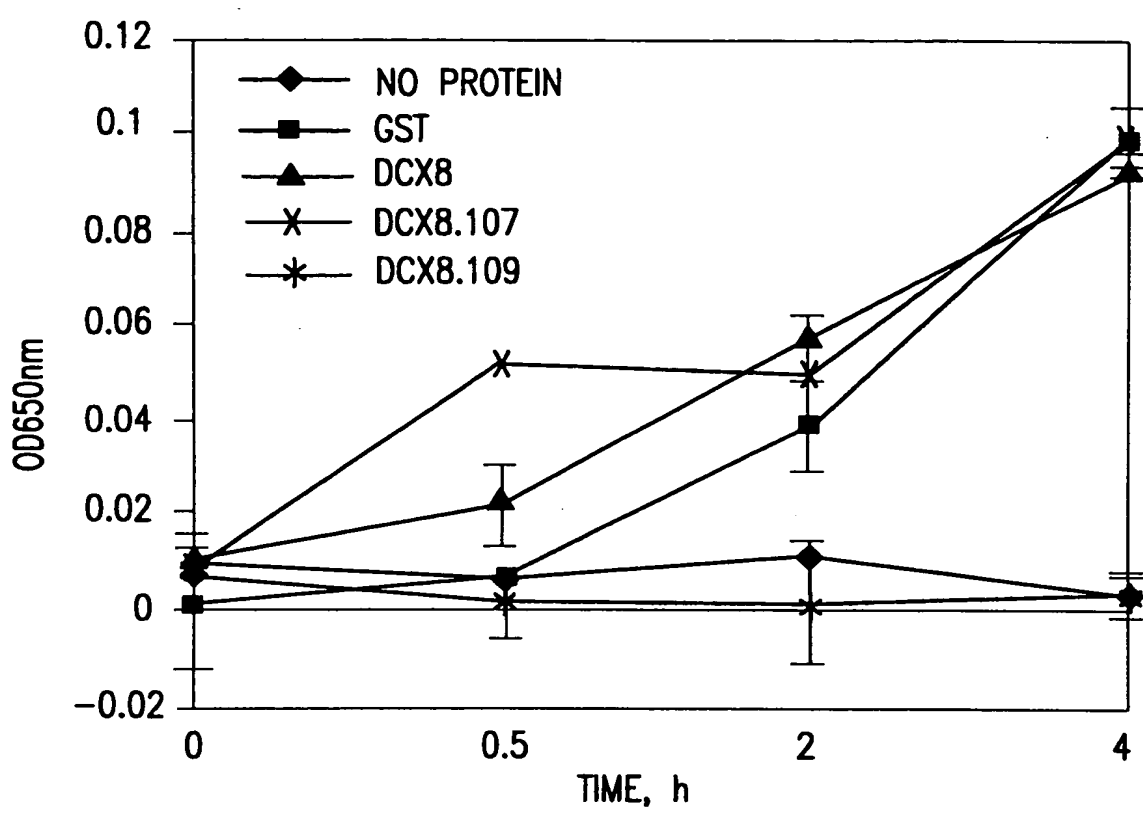


FIG.8C

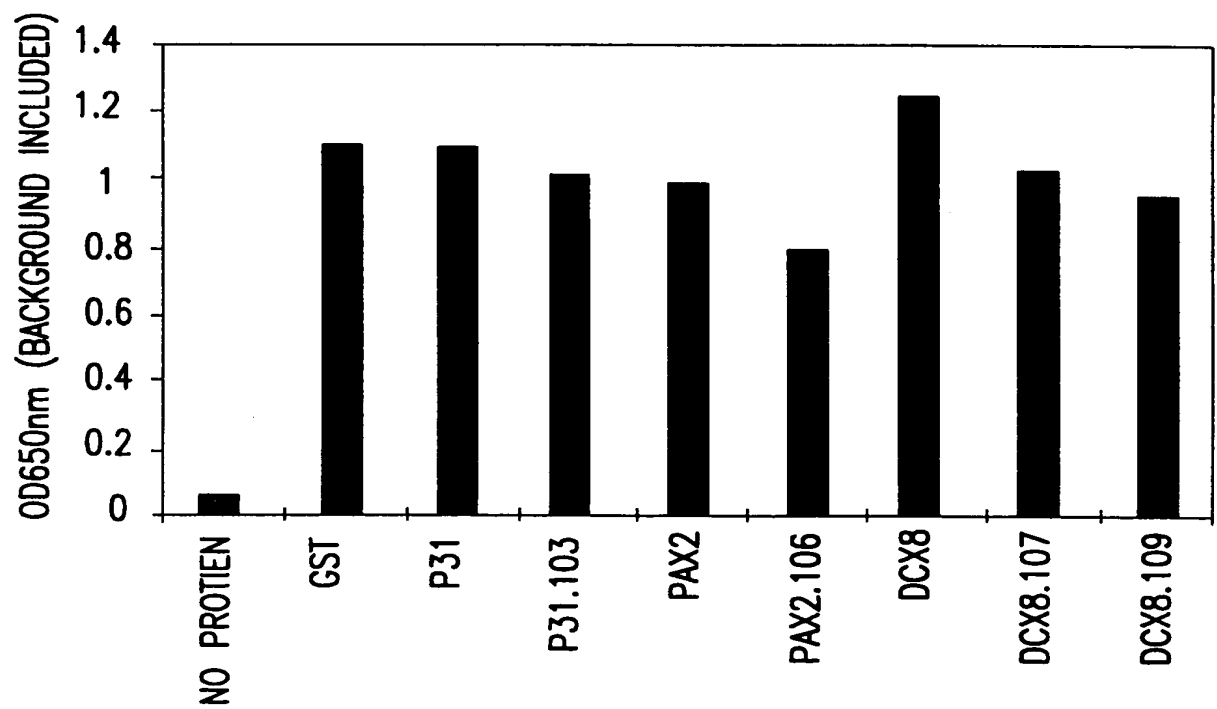


FIG.8D

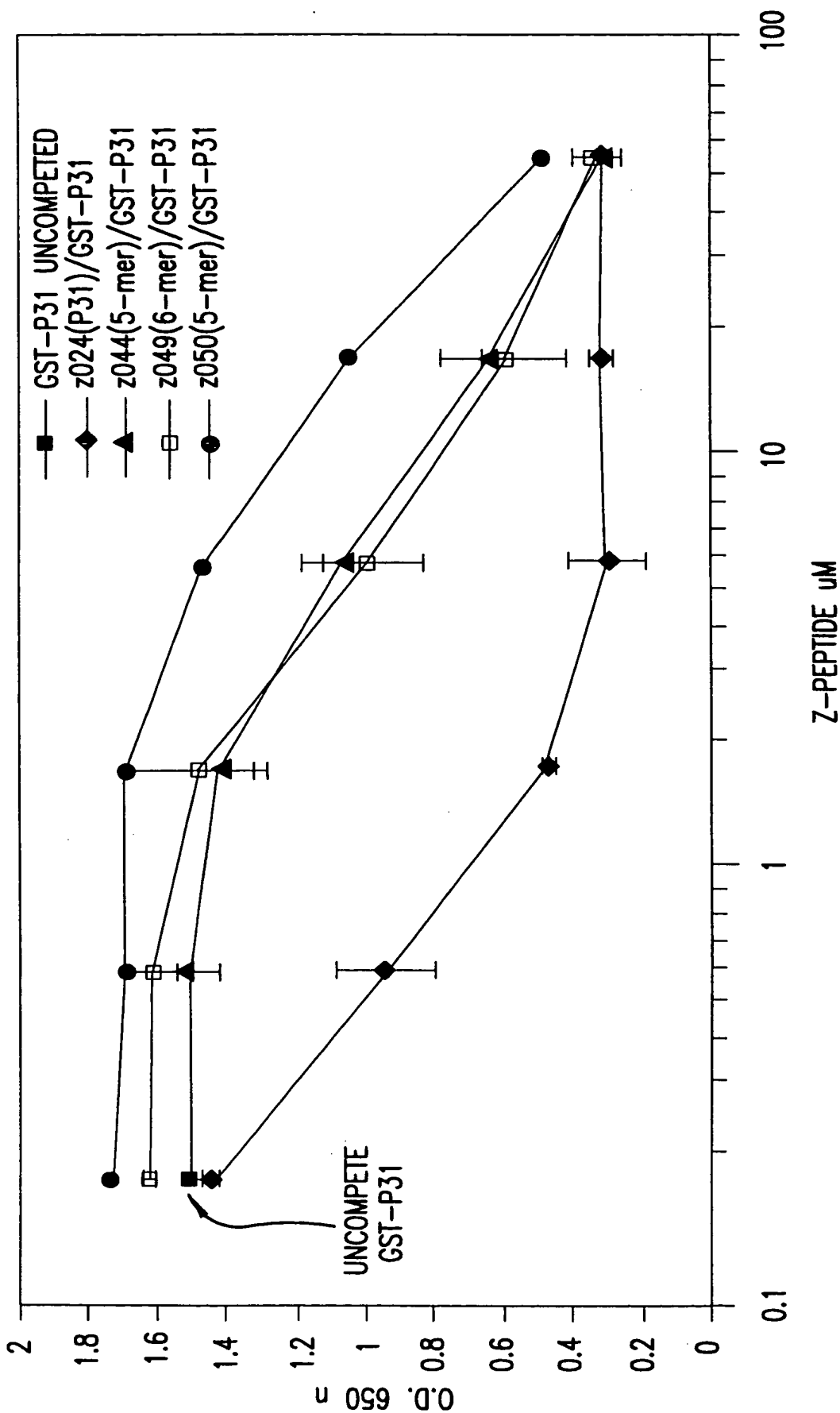


FIG.9A

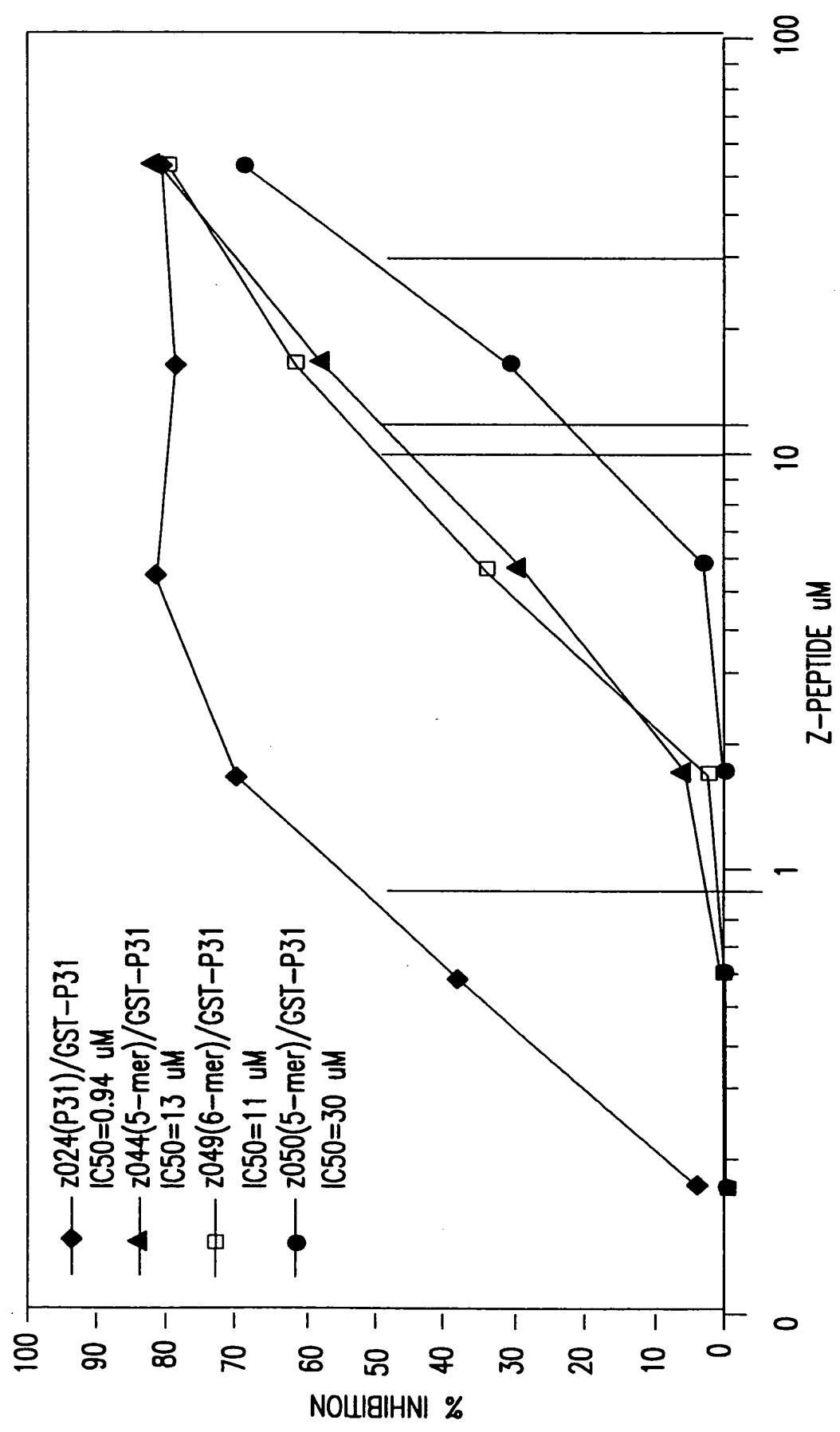


FIG.9B

Peptide Name	Sequence	pI	IC ₅₀	GST/C2BBel
ELAND24 (P31)	1 10 20 30 40 SARDSGPAEDGSRVRLNGVENANTRKSSRSNPRGRHPG SARDSGPAEDGSRVRLNG	11.88	0.5-2.2	+++
101	DGSRVRLNGVENANTRKSSR			-
102	ENANTRKSSRSNPRGRHP			++
103	ENANTRKSSR			-
110	RKSSRSNPRG			-
111	SNPRGRKRP			-
112	TRKSSRSNPRG			
119	ZENANTRKSSRSNPRGRHPG	12.28	0.5-1.7	
Z28	ZTRKSSRSNPRGRHPG	12.40	5.5-15	
Z29	ZENANTRKSSRSNPRG	11.81	> 50	
Z30	ZTRKSSRSNPRGRHPG	12.70	0.6-3.2	
Z31	ZENANTRKSSR	10.89	> 50	
Z39	ZSNPRGRHPG	12.40	5.9-29	
Z40	ZENANT	3.75	>50	
Z41	ZANTRKS	11.05	>50	
Z42	ZTRKSS	11.05	>50	
Z43	ZRKSSR	12.11	13- > 50	
Z44	ZKSSRSN	11.05	40-48	
Z45	ZSSRSNPG	10.04	>50	
Z46	ZRSNPRG	12.40	>50	
Z47	ZSNPRG	10.04	>50	
Z48	ZPRGRRH	12.40	11-20	
Z49	ZRRHPG	12.10	30	
Z50	ZKSSRGN	12.40	>50	
Z51 (HepC core)	ZKSSRGN	12.10	9.8	
Z52 (HepC p26664)	ZKTSERSQPRGRHPG			
Z53	ZTrKSSrSNPrGrHPG			1.6
Z54	ZTRKSSrSNPRGrHPG			1.6
Z21 (HAX42)	SDHALGTNLPSDNAKEPGDYNCCGNGNSTGRKVFNRRPISAIP	11.27	1.7	

FIG.10A

Peptide Name	Sequence	1	10	20	30	40	pI	IC ₅₀	GI/C28Be1
ELANO18(PAX2)	STPPSREAYSRPYSDSDTNAKHSSHNRLRTRSPNG						10.88	0.6-0.9	1
104	STPPSREAYSRPYSDSDSD							+++	
105	SRPYSDSDSDTNAKHSSHN								-
106	TNAKHSSHNRLRTRSPN								++
113	TNAKHSSHN								-
114	SSHNRLRTR								+/-
115	RRLRTRSPN								+/-
Z32	ZTNAKHSSHNRLRTRSPN						12.7	1.2	
Z33	ZTNAKHSSHNRLRTR						12.58	1.6	
Z34	ZSSHNRLRTRSPN						12.7	1.6	1.3
Z35	ZSSHNRLRTR						12.58	0.38 - 1.8	0.68, 1.5
Z26	ZSEANLDGRKSRYPSPRRNSSTRPSPNSVHARYPSTDHD						10.88	7-8	3
Z38	ZSRANTDGRKSRYPSPRRNSSTEPRLSPNSVHARYPSTDHD						10.88	1.7	0.9
Z55	ZTNAKHSSHN							42	
Z56	ZRRLRTRSPN							1.7	
Z57	ZRRLRTRSP							1.9	
Z58	ZRRLRTR							3.4	
Z59	ZrLrTrSPN							NOT DONE	
Z73	ZASHNRLRTR						1.5	5.5	
Z74	ZSAHNRLRTR						6.2		
Z75	ZSSANRLRTR						1.6		
Z76	ZSSHARLRT						1.8		
Z77	ZSSHNARLRT						3.9	5.2	
Z78	ZSSHNARLRT						4.5	4.6	
Z79	ZSSHNRRARTR						1.4		
Z80	ZSSHNRLRATR						3.4	5.2	
Z81	ZSSHNRLRAR						2.2		
Z82	ZSSHNRLRTA						3.4		
Z21 (HAX42)	ZSDHALGTNLRSDNAKEPGDYNCCGNGNSTGRKVFNRRRPSAIP						11.27	0.7	

FIG.10B

SN: 10	Peptide Name	Sequence	1	10	20	30	40	pI	IC ₅₀	GST/C28BeI
	ELANO16 (SN110)	RVGQCTSDVRRPWARSCAHQGGAGTRNSHGCI	1	10	20	30	40			
	116	RVGQCTSDVRRPWARSCA						10.19	0.22	++
	117	VRRPWARSCAHQGGAGTRNS								-
	118	GTRNSHGCI								+
	Z17	ZRVGQCTSDVRRPWARSCAH						8.66	3.6	+/-
	Z16C23	ZCGAGTRNSHGCI						9.03	0.7	
	Z36	ZVRRPWARSCAHQGGAGTRNS						11.62	0.27	
	Z37	ZCTDSDVRRPWARSC						8.01	3	
	Peptide Name	Sequence	1	10	20	30	40	pI	IC ₅₀	GST/C28BeI
	ELANO21 (HAX42)	SDHALGTNLRSDNAKEPGDYNCNGNSTGRKVFNR	1	10	20	30	40			
	ELANO18 (PAX2)	STPPSREAYSRPYSVDSDTNAKHSSHNRLR						11.27	5.5	++
	Z26	ZSEANLDGRKSRYSPPRNSSTRPRTSPNSVHARYP						10.88	0.23	+++
	Z38	ZSRANTDGRKSRYSPPRNSSTEPRLSPNSVHARYP						10.88	<0.2	
	Z34 (PAX2 14mer)	ZSSHNRRLRTRSRPN						12.7	0.33	

FIG.10C

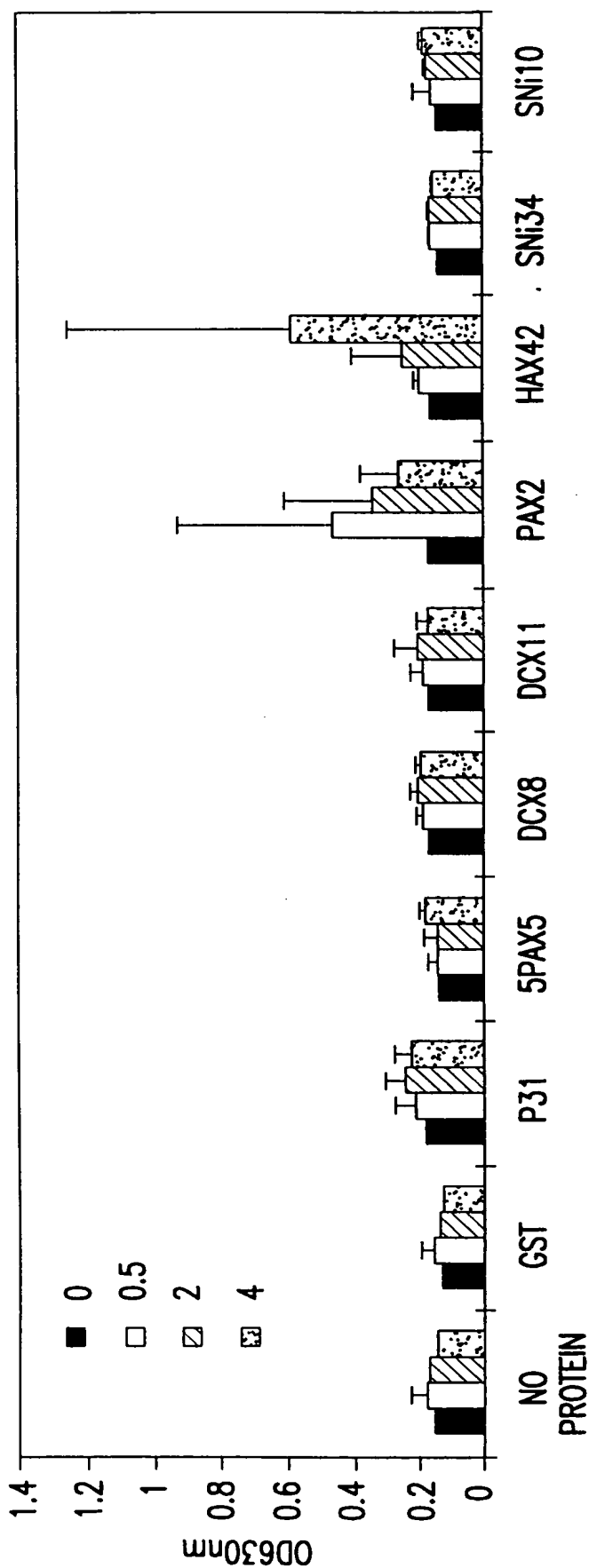


FIG.11A

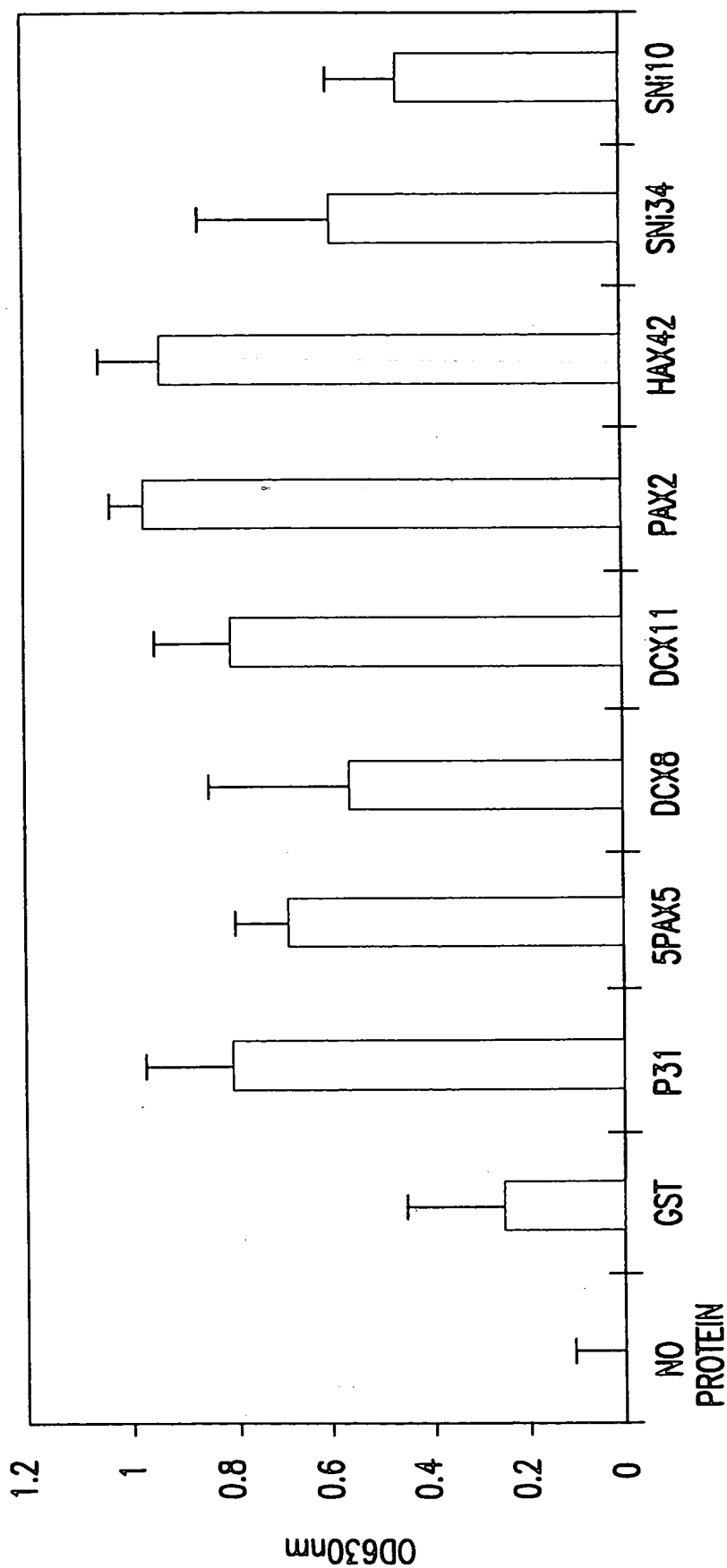


FIG. 11B

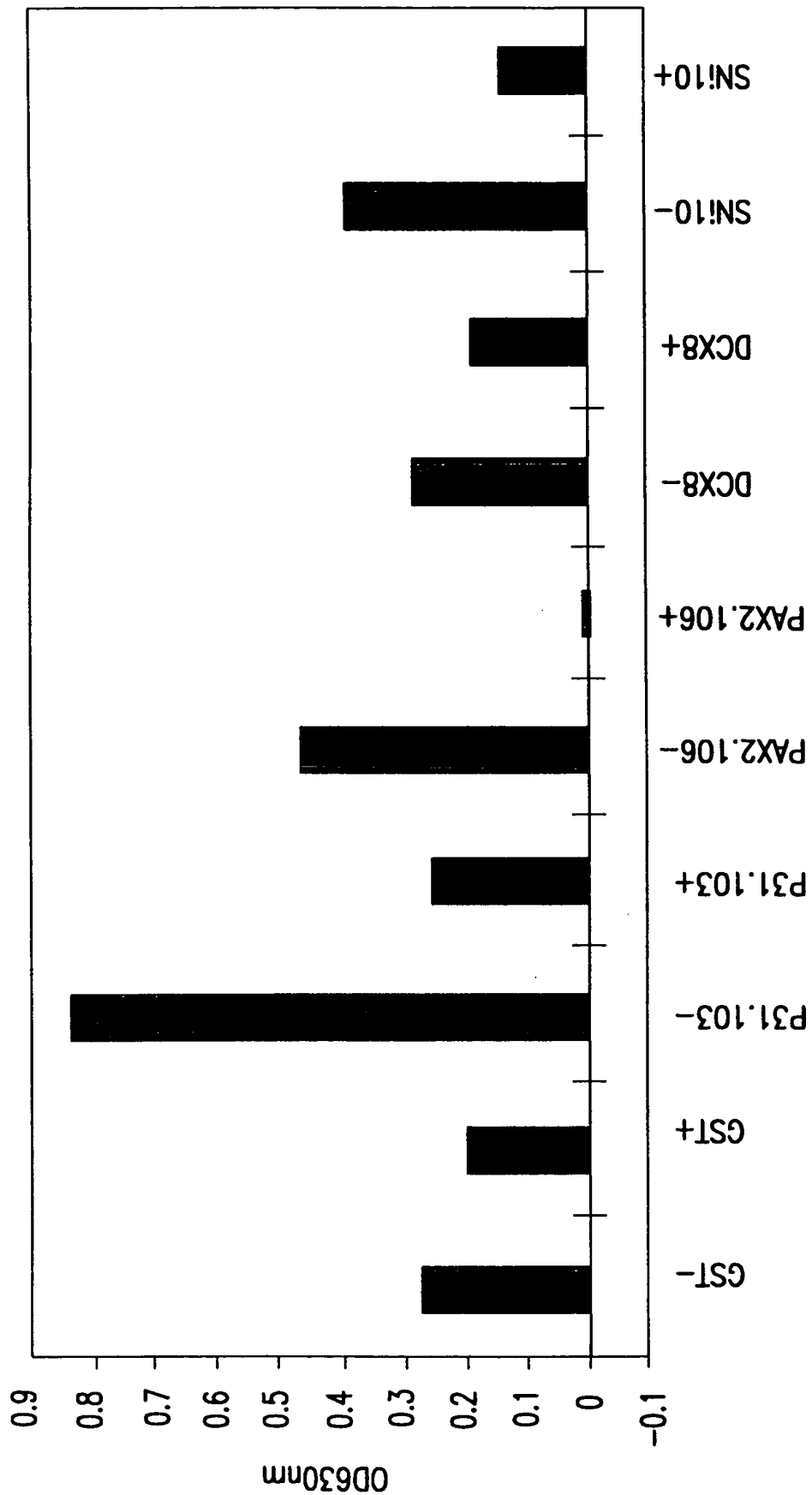


FIG.12

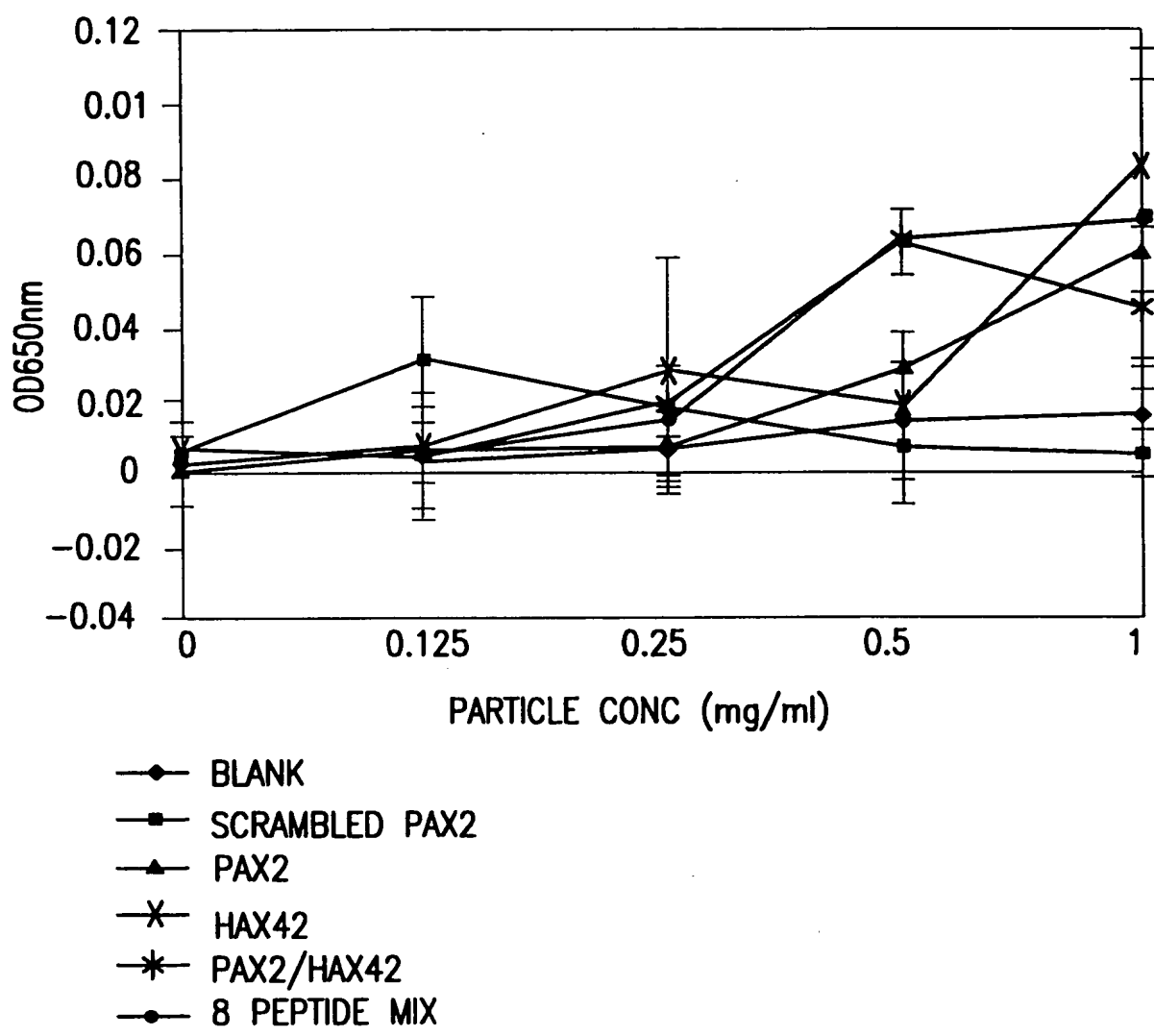


FIG.13A

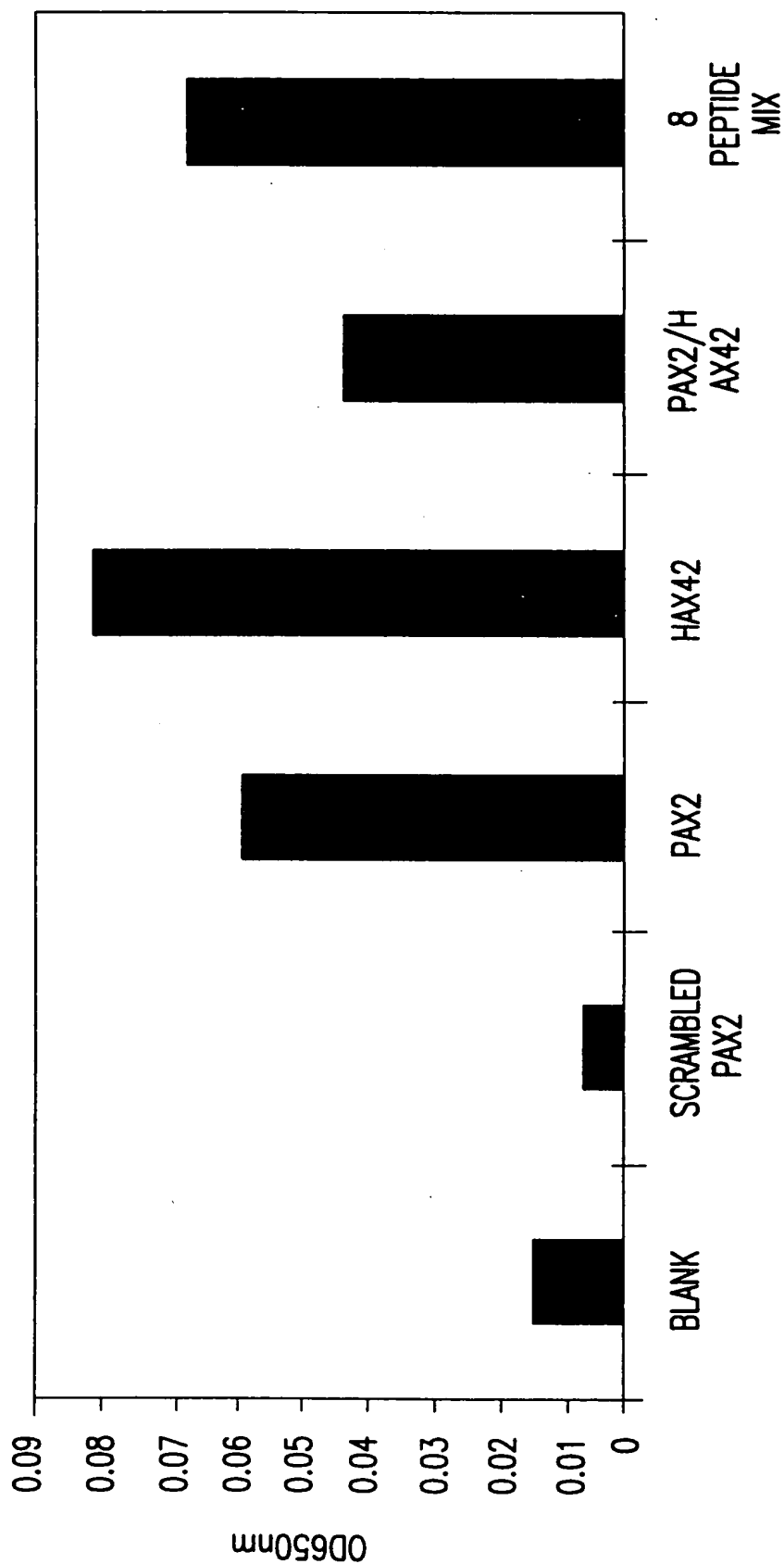


FIG.13B

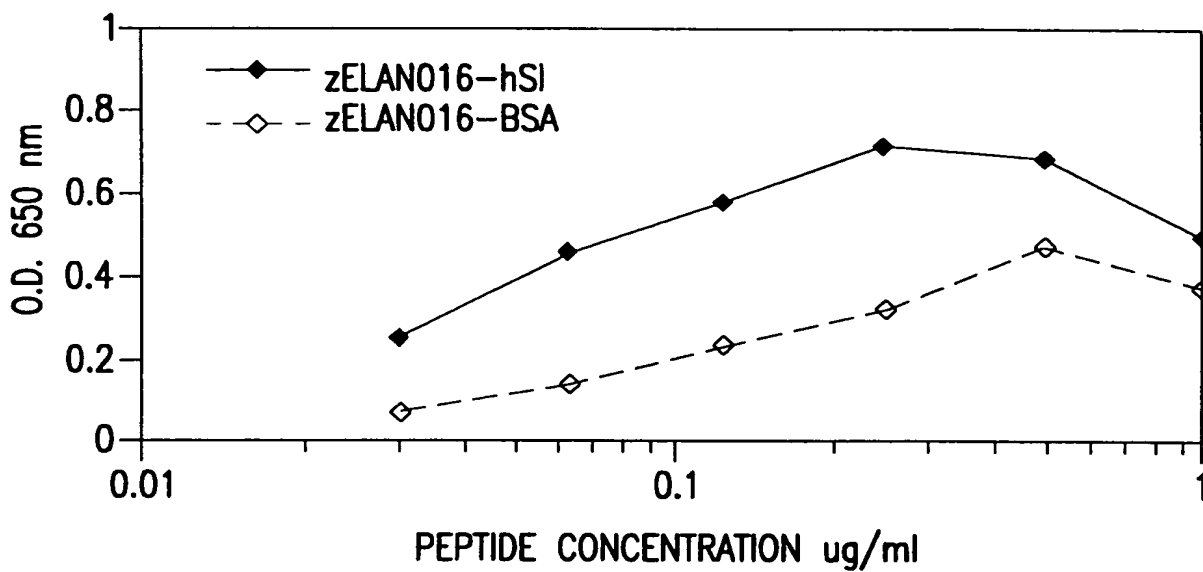
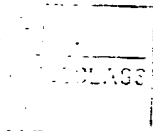


FIG. 14A

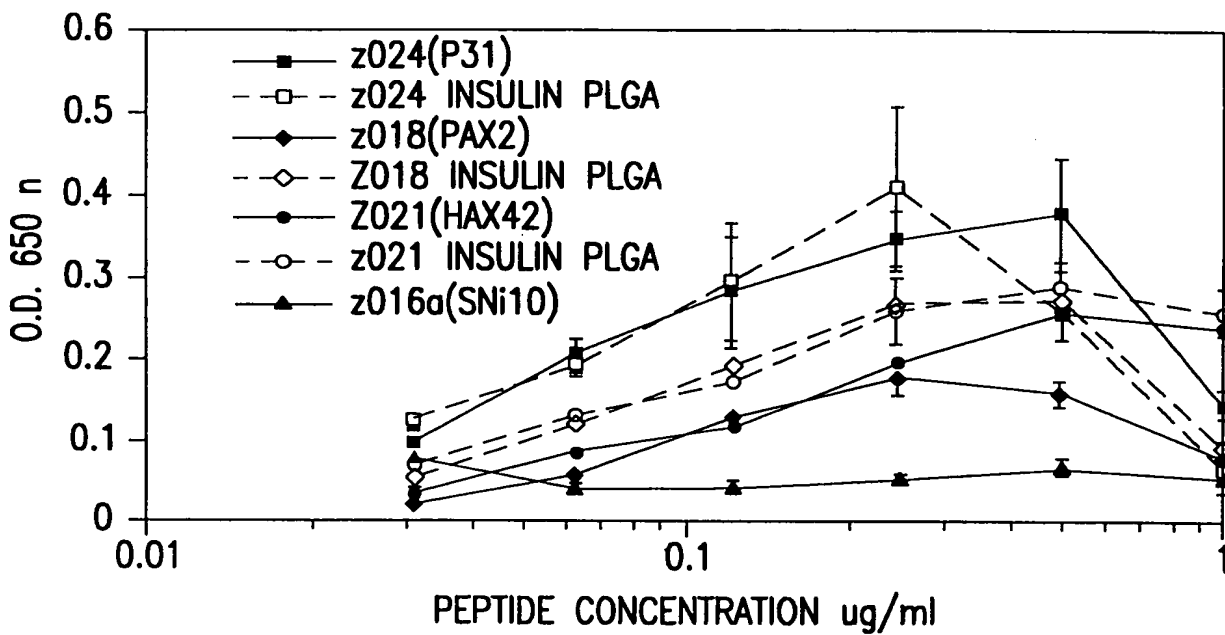


FIG. 14B

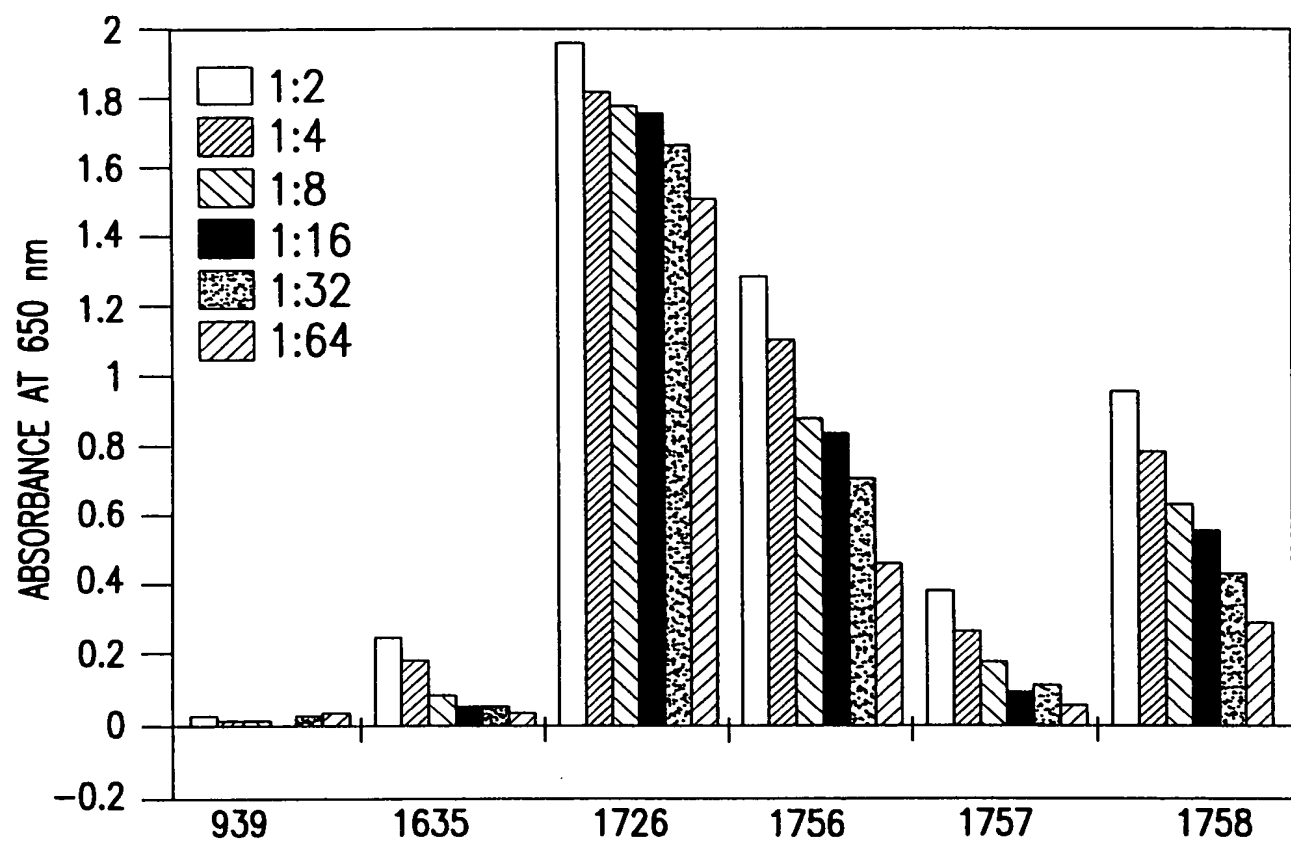


FIG. 15A

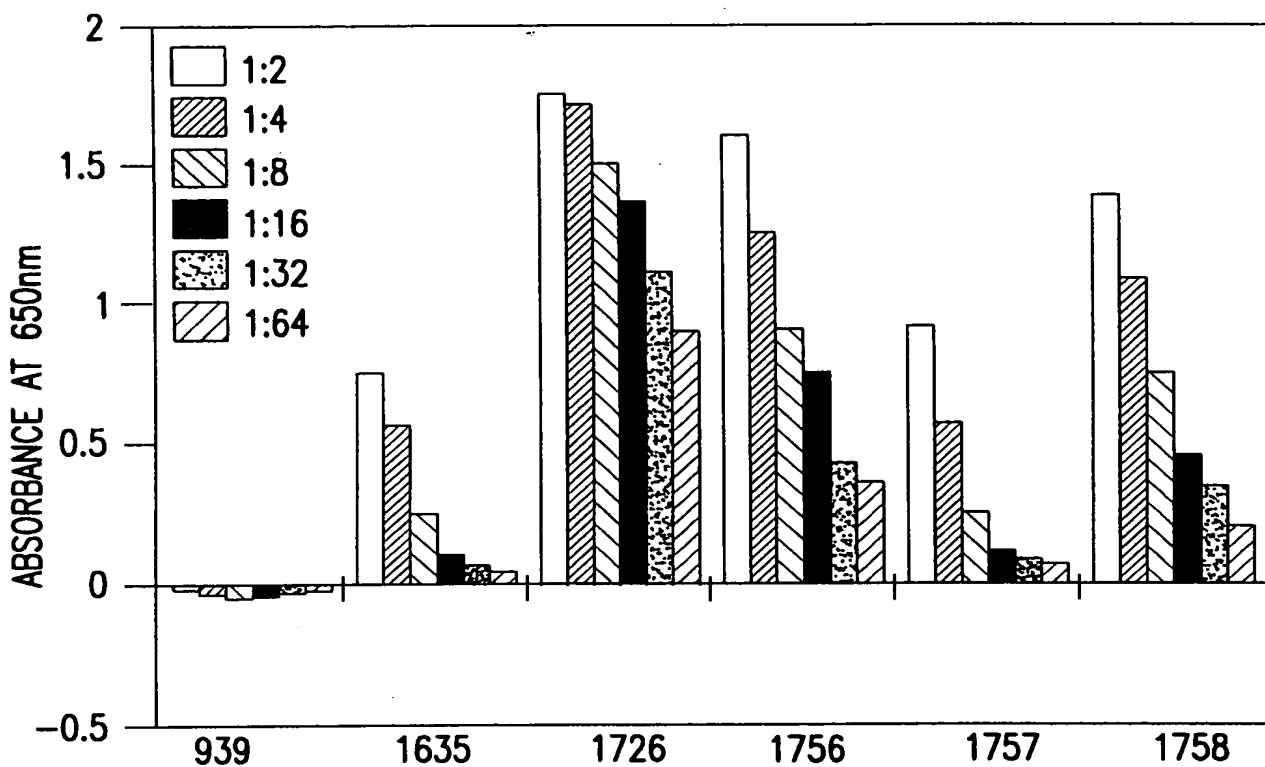


FIG. 15B

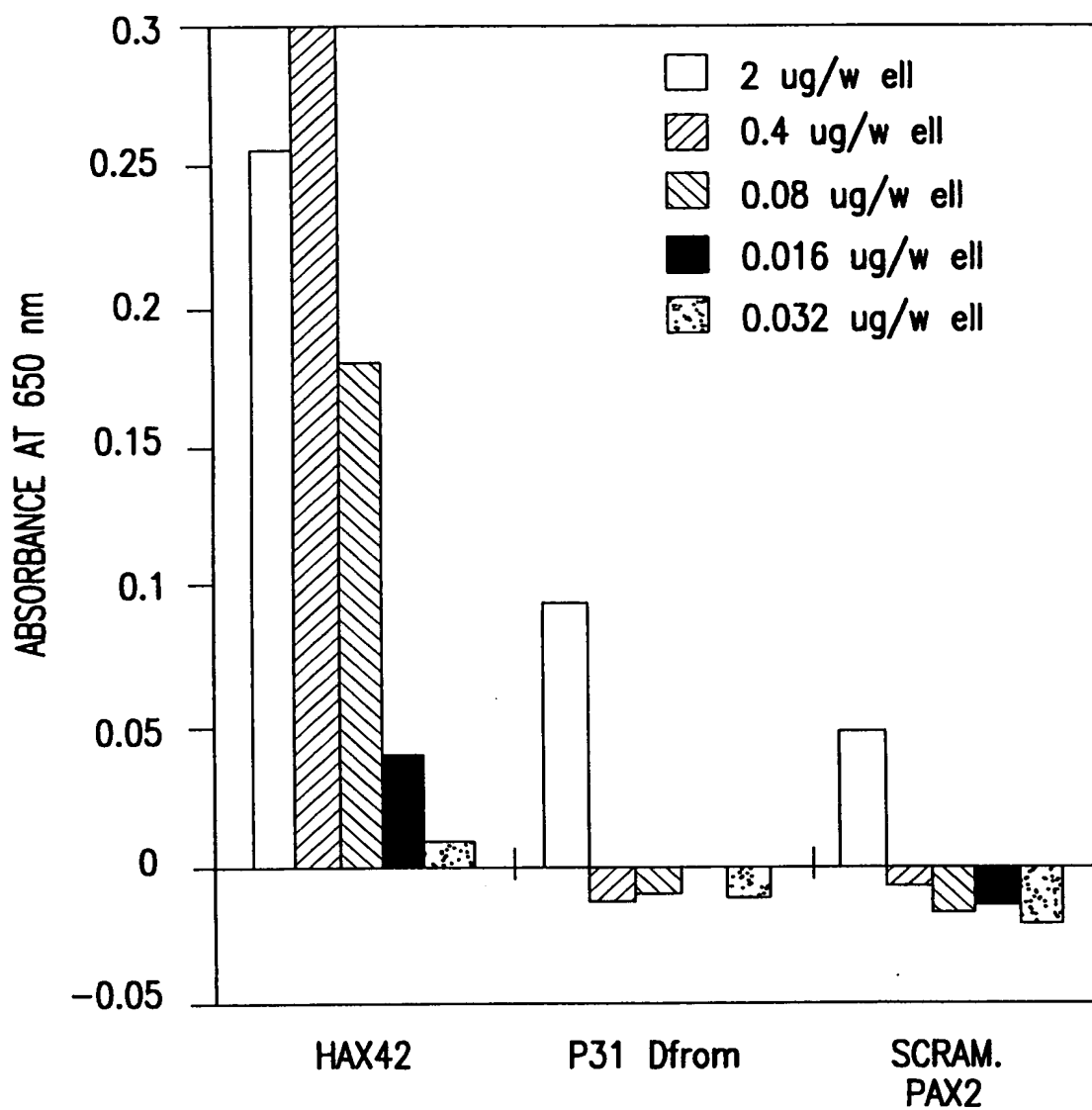


FIG.16A

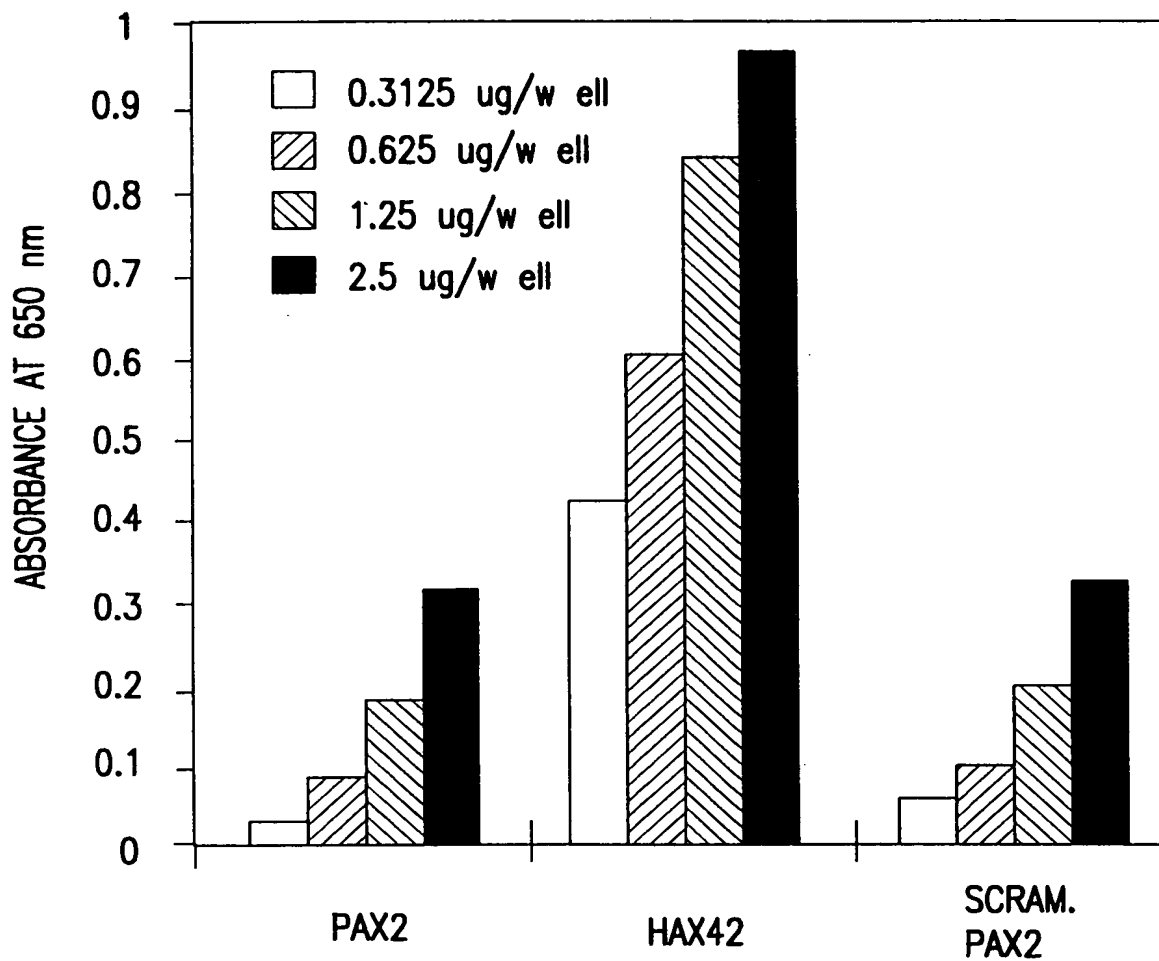


FIG.16B

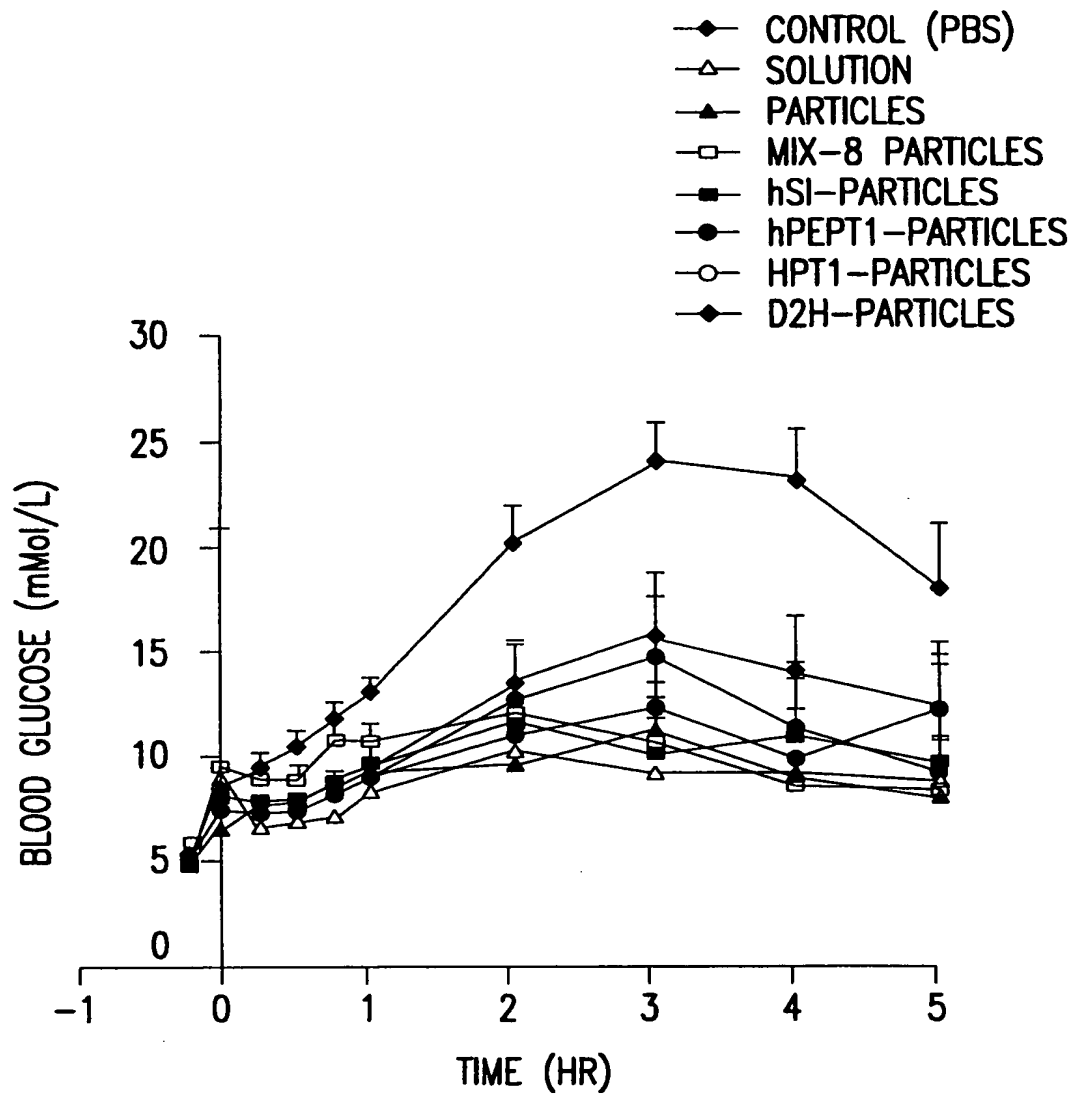


FIG. 17A

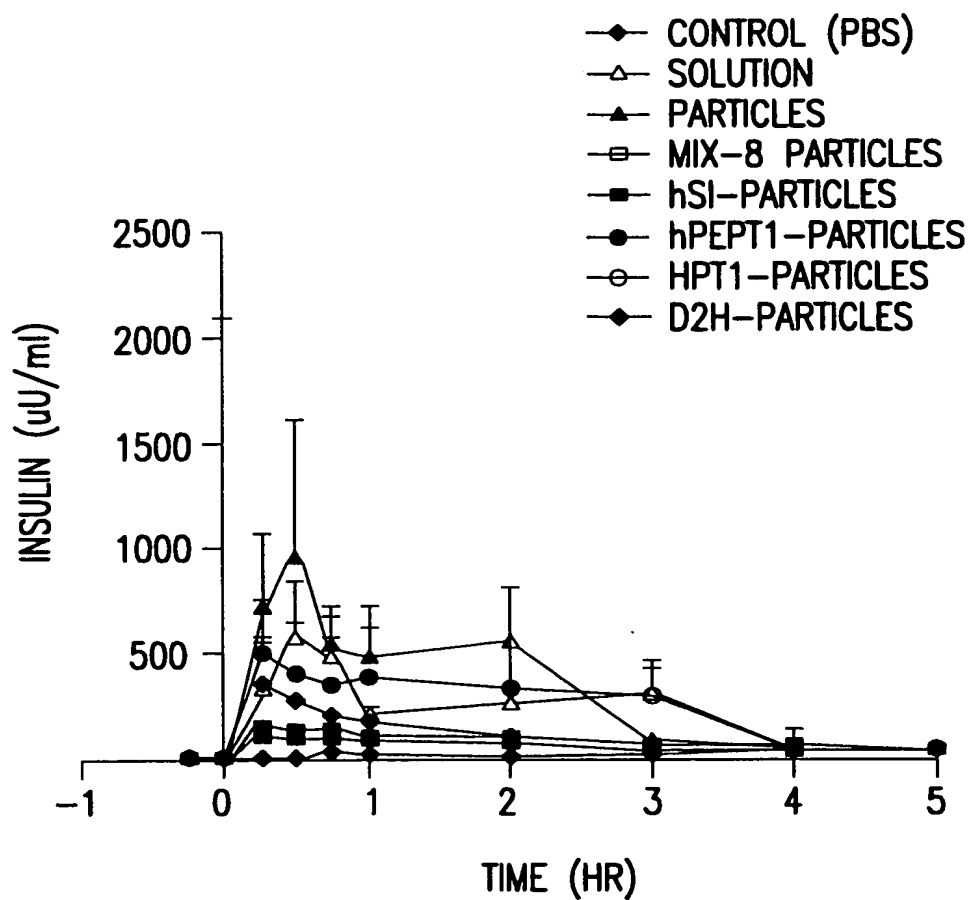


FIG. 17B

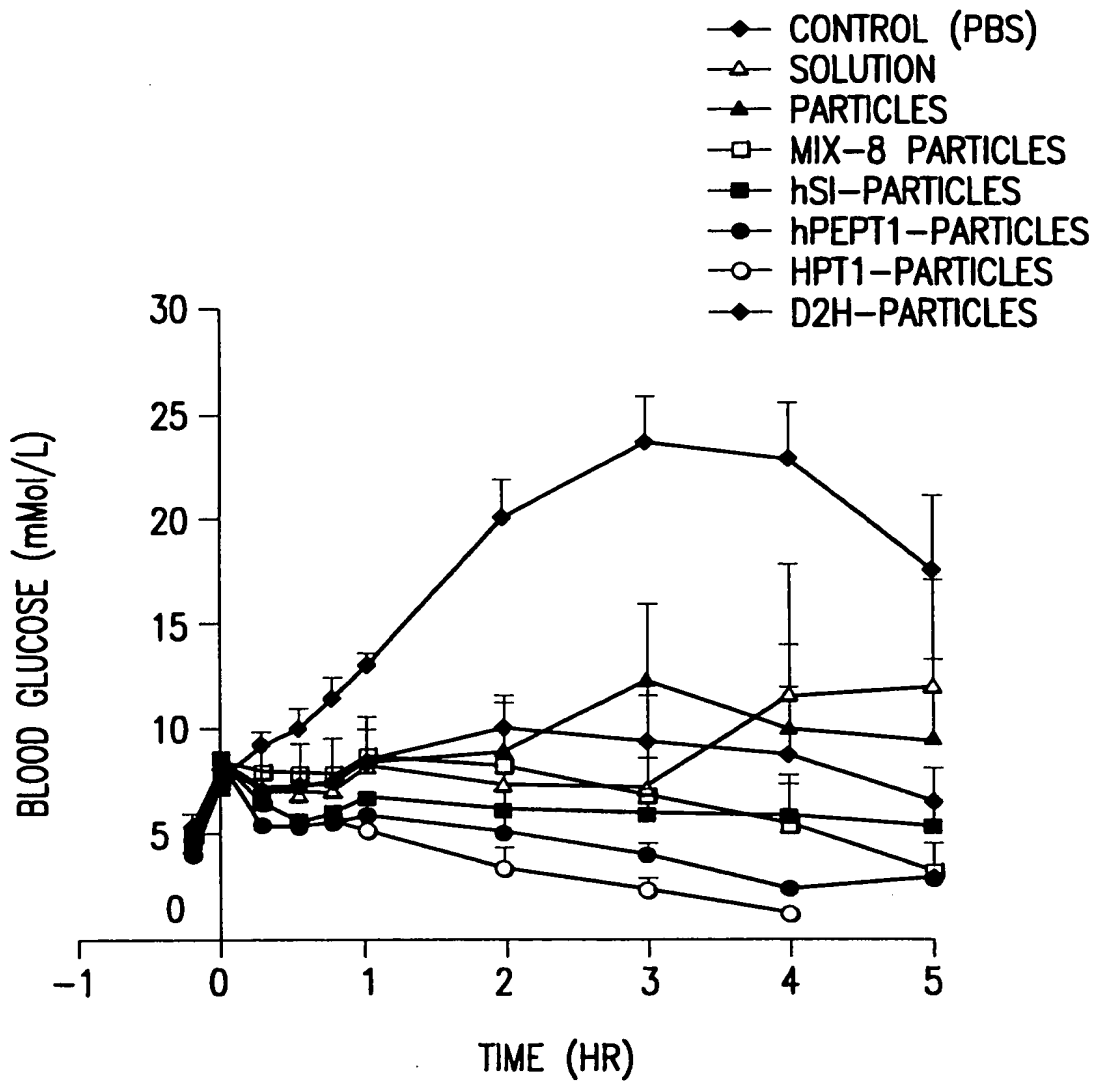


FIG. 18A

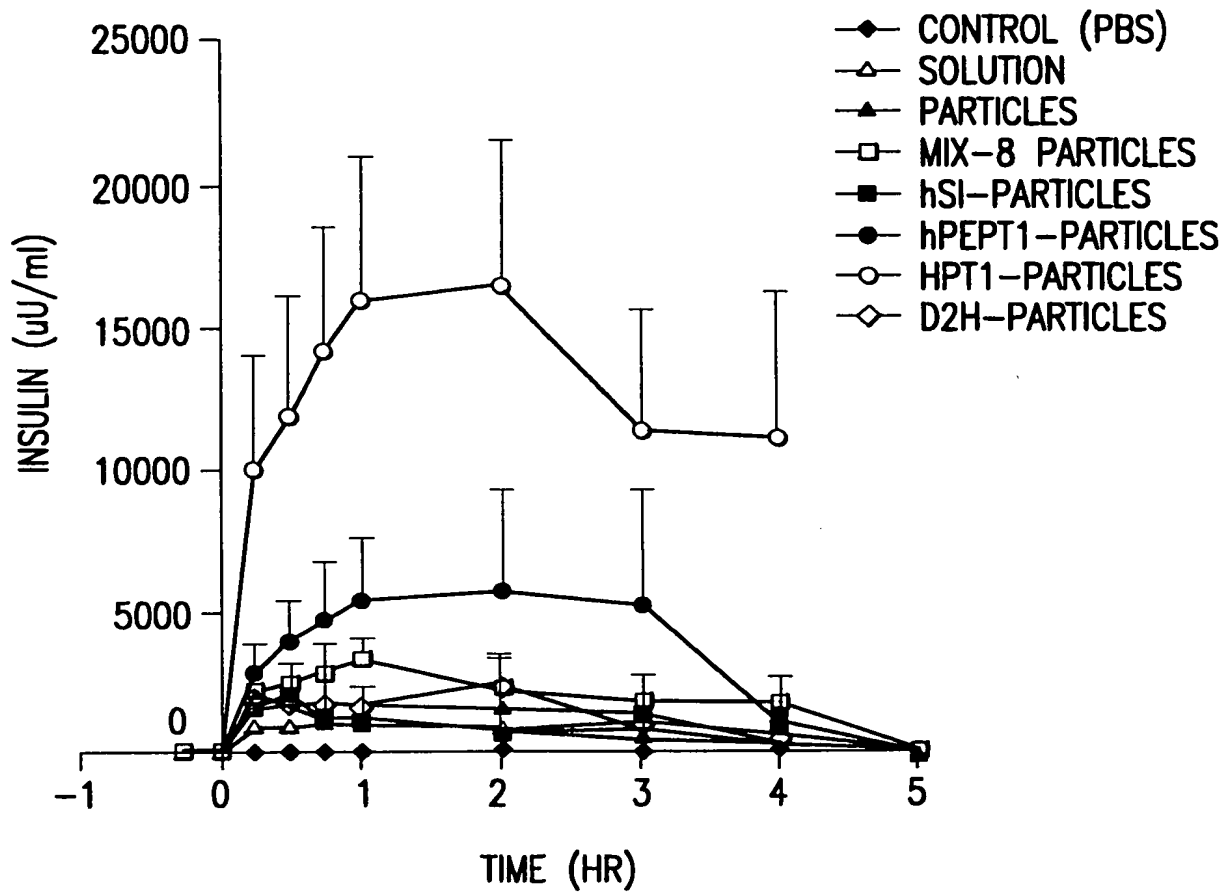


FIG.18B

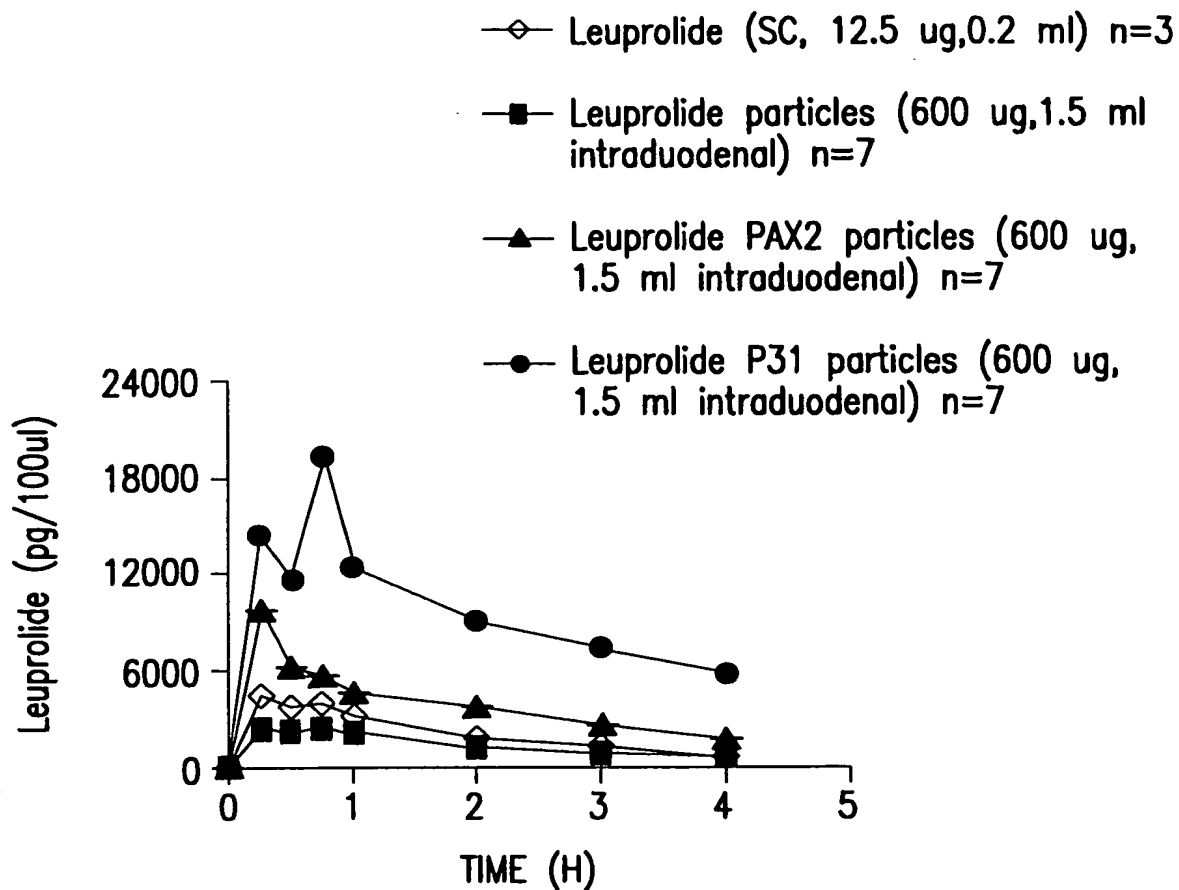


FIG. 19

P31 AA SEQ. POSITION	KNOWN PROTEIN	HOMOLOGOUS SEQ. POSITION
12-34	FASCICULIN 2	10-32
4-12	MESENTERICOPEPTIDASE	54-62
15-31		175-191
26-39	CORE PROTEIN (HEPATITIS C VIRUS)	5-18
26-39		11-24
26-39		21-34
26-39		38-51
23-30		39-55
25-39		41-55
26-39		51-64
16-39	PT-NANBH POLYPROTEIN N-TERMINUS	51-64
28-40	AL2 PROTEIN (CAENORHABDITISELEGANS)	70-82
26-38	CAPSID PROTEIN (HEPATITIS C VIRUS TYPE 3g)	48-60
26-39	GENOME POLYPROTEIN (HEPATITIS C VIRUS)	57-70

FIG.20

DCX8AA SEQ. POSITION	KNOWN PROTEIN	HOMOLOGOUS SEQ. POSITION
20-27	ENDO-1,4-BETA-D-GLUCANASE	78-85
30-37		221-228
21-34	P-HYDROXYBENZOATE HYDROXYLASE	285-298
5-15		54-64
7-21	CYTOCHROME	50-64
7-21	CYTOCHROME C3	50-64
	TRIMETHYLARNINE DEHYDROGENASE	208-219
32-43		396-407
30-37	Gag-JunD FUSION PROTEIN	24-31
26-30		16-20
23-44	SECRETIN PRECURSOR, N-PROSECRETIN, SECRETIN ANIDE	18-39
33-44	T-CELL RECEPTOR V BETA CHAIN	15-26
27-33		3-9
23-44	SECRETIN PRECURSOR PIR	18-39
31-44	HYPOTHETICAL PROTEIN V (SYNECHOCYSTIS)	275-288
24-30		251-257
23-43	PUTATIVE RNA BINDING PROTEIN	230-250
28-40	Mu SON OF SEVENLESS 1	1-13
24-35	NEUROPEPTIDE PRECURSOR	80-91
29-43		5-19
23-43	RNA-BINDING PROTEIN (MACACAFASCICULARIS)	230-250
23-43	RNA-BINDING PROTEIN (HOMOSAPIENS)	230-250
23-43	AUTOSOMAL GENE-AZOOSPERMIA FACTOR	230-250
25-38	COLLAGEN	25-28
24-35		4-15
29-41	PROBABLE CELL GROWTH REGULATOR	306-318
24-35	RIBOSOMAL PROTEIN S2	24-35
T6-39		182-185
24-44	CAENORHABDITIS ELEGANS	296-316
23-34	pid:e208155 (HOMO SAPIENS)	61-72
36-43		116-123

FIG.21A

DCX8A SEQ. POSITION	KNOWN PROTEIN	HOMOLOGOUS SEQ. POSITION
24-38	XYLULOSE KINASE	16-30
24-39	CAENORHABDITIS ELEGANS	57-72
26-42		65-81
27-33	HYPOTHETICAL PROTEIN-PHAGE BZ13	22-28
35-39		31-35
30-42	CEREBELLIN-LIKE GLYCOPROTEIN	2-14
8-22	DNA PRIMASE	170-184
2-7		76-81
5-21	COAT PROTEIN (BEAN COMMON MOSAIC VIRUS)	12-28
5-21	COAT PROTEIN (BEAN COMMON MOSAIC VIRUS)	33-49
5-21		19-35
5-21	POLYPROTEIN (BEAN COMMON MOSAIC VIRUS)	215-231
5-21		39-55
5-21	Nib PROTEINLCOAT PROTEIN (COWPEA APHID-BOME MOSAIC VIRUS)	92-108
2-13	MHC CLASS 1 PIPI (PITHECIA)	111-122
14-22		236-334
3-19	TALIN (CAENORHABDITIS ELEGANS)	1538-1554
2-9	ACETAMIDASE PIR	359-366
9-20		483-494
10-16	RHIZOBIONS ETLI STRAIN	134-140
17-30		173-186
31-39		200-208
2-11	NEUROTOXIN 1 (TOXIN B) A. STOKESI	7-16
12-33		26-47
21-27	SUID HERPES VIRUS 1 EARLY PROTEIN	425-432
30-43		51-64
13-42	RICE cDNA PARTIAL SEQUENCE	50-151
8-15	FUSION PROTEIN	24-31
4-8		16-20
1-22	SECRETIN PRECURSOR, N-PROSECRETIN, SECRETIN-AMIDE	18-39
11-22	T-CELL RECEPTOR V BETA CHAIN	15-26
5-11		3-9
9-22	HYPOTHETICAL PROTEIN	275-288
2-8		251-257

FIG.21B

DCX8A SEQ. POSITION	KNOWN PROTEIN	HOMOLOGOUS SEQ. POSITION
1-21	PUTATIVE RNA BINDING PROTEIN	230-250
6-18	HYPOTHETICAL PROTEIN-MOUSE PIR	1-13
2-13	NEUROPEPTIDE PRECURSOR	80-91
7-21	orf3-HUMAN	5-19
1-21	RNA-BINDING PROTEIN	230-250
13-16	COLLAGEN	25-28
7-19	PROBABLE CELL GROWTH OR DIFFERENTIATION REGULATOR	306-318
2-13	RIBOSOMAL PROTEIN S2	14-25
14-17		182-185
2-22	CAENORHABDITIS ELEGANS	296-316
1-12	HOMOSAPIENS	61-72
14-21		116-123
2-16	XYLULOSE KINASE	16-30
8-15	T CELL RECEPTOR DELTA CHAIN	55-62
5-8		12-15
8-17	SEQ. 43 FROM PATENT US	12-21

FIG.21C

CLASS

DRAFTS

DAB10 AA SEQ. POSITION	KNOWN PROTEIN	HOMOLOGOUS SEQ. POSITION
13-34	1,3-BETA-GLUCANASE	231-252
3-11	PHOTOSYNTHETIC REACTION CENTER	20-28
16-27		128-139
28-35	MYB PROTO-ONCOGENE PROTEIN	131-138
5-18		32-45
23-36	LYSOZYME MUTANT	130-143
28-35	LIPASE	400-407
3-15		159-171
3-37	TRYPSIN	169-203
13-34	1,3-1,4-BETA-GLUCANASE	232-253
4-10	LACTATE DEHYDROGENASE	190-196
11-7		244-250
4-10	APO-LACTATE DEHYDROGENASE	190-196
11-17		244-250
4-10	LACTATE DEHYDROGENASE	191-197
11-17		245-251
16-26	OVOTRANSFERRIN	240-250
23-36	GENOME POLYPROTEIN MATRIX PROTEIN	1022-1035
14-20	ROUS SARCOMA VIRUS	43-49
2-12		13-23
14-20	HYPOTHETICAL PROTEIN-AVIAN LEUKOSIS VIRUS	43-49
4-20	T CELL RECEPTOR DELTA CHAIN VARIABLE REGION	1-4
14-18		12-16
2-12	GAG POLYPROTEIN-AVIAN ENDOGENOUS VIRUS RAV-0	139-149
14-20		169-175
	p19 PROTEIN-AVIAN ERYTHROBLASTOSIS VIRUS	189-199
14-20		219-225
7-19	ALI PROTEIN-POTATO YELLOW MOSAIC VIRUS	222-234
3-22	ENDO-1,4-BETA GLUCANASE	186-205
6-18	I _o PROTEIN-BROME MOSAIC VIRUS	430-442
2-12	GAG POLYPROTEIN-FUJINAMI SARCOMA VIRUS	186-196
14-22		216-222
2-12	GAG PROTEIN-ROUS SARCOMA VIRUS	190-200
14-20		220-226
1-12	CORTICOTROPIN-LIKE INTERMEDIATE LOBE PEPTIDE	7-18
1-22	GENE PRODUCT (CAENORHABDITIS ELEGANS)	4-25
31-37	T CELL RECEPTOR DELTA CHAIN	56-62
26-39		12-15
26-37	LYSOZYME MUTANT	133-144

FIG.22

AT 101
 CLASS
 DRAFT

ATG Met 1	TCC Ser	CCT Pro	ATA Ile	CTA Leu 5	GGT Gly	TAT Tyr	TGG Trp	AAA Lys	ATT Ile 10	AAG Lys	GGC Gly	CTT Leu	GTG Val	CAA Gln 15	CCC Pro	48
ACT Thr	CGA Arg	CTT Leu	CTT Leu 20	TTG Leu	GAA Glu	TAT Tyr	CTT Leu	GAA Glu 25	GAA Glu	AAA Lys	TAT Tyr	GAA Glu 30	GAG Glu	CAT His	TTG Leu	96
TAT Tyr	GAG Glu	CGC Arg 35	GAT Asp	GAA Glu	GGT Gly	GAT Asp 40	AAA Lys	TGG Trp	CGA Arg	AAC Asn	AAA Lys 45	AAG Lys	TTT Phe	GAA Glu	TTG Leu	144
GGT Gly 50	TTG Leu	GAG Glu	TTT Phe	CCC Pro	AAT Asn	CTT Leu 55	CCT Pro	TAT Tyr	TAT Tyr	ATT Ile	GAT Asp 60	GGT Gly	GAT Asp	GTT Val	AAA Lys	192
TTA Leu 65	ACA Thr	CAG Gln	TCT Ser	ATG Met	GCC Ala 70	ATC Ile	ATA Ile	CGT Arg	TAT Tyr	ATA Ile 75	GCT Ala	GAC Asp	AAG Lys	CAC His	AAC Asn 80	240
ATG Met	TTG Leu	GGT Gly	GGT Gly 85	TGT Cys	CCA Pro	AAA Lys	GAG Glu	CGT Arg	GCA Ala 90	GAG Glu	ATT Ile	TCA Ser	ATG Met	CTT Leu 95	GAA Glu	288
GGA Gly	GCG Ala	GTT Val	TTG Leu 100	GAT Asp	ATT Ile	AGA Arg	TAC Tyr	GGT Gly 105	GTT Val	TCG Ser	AGA Arg	ATT Ile	GCA Ala 110	TAT Tyr	AGT Ser	336
AAA Lys	GAC Asp	TTT Phe 115	GAA Glu	ACT Thr	CTC Leu	AAA Lys 120	GTT Val	GAT Asp	TTT Phe	CTT Leu	AGC Ser	AAG Lys 125	CTA Leu	CCT Pro	GAA Glu	384
ATG Met 130	CTG Leu	AAA Lys	ATG Met	TTC Phe	GAA Glu	GAT Asp 135	CGT Arg	TTA Leu	TGT Cys	CAT His	AAA Lys 140	ACA Thr	TAT Tyr	TTA Leu	AAT Asn	432
GGT Gly 145	GAT Asp	CAT His	GTA Val	ACC Thr	CAT His 150	CCT Pro	GAC Asp	TTC Phe	ATG Met	TTG Leu 155	TAT Tyr	GAC Asp	GCT Ala	CTT Leu	GAT Asp 160	480
GTT Val	GTT Val	TTA Leu	TAC Tyr	ATG Met 165	GAC Asp	CCA Pro	ATG Met	TGC Cys	CTG Leu 170	GAT Asp	GCG Ala	TTC Phe	CCA Pro	AAA Lys 175	TTA Leu	528
GTT Val	TGT Cys	TTT Phe	AAA Lys 180	AAA Lys	CGT Arg	ATT Ile	GAA Glu	GCT Ala 185	ATC Ile	CCA Pro	CAA Gln	ATT Ile	GAT Asp 190	AAG Lys	TAC Tyr	576
TTG Leu	AAA Lys	TCC Ser 195	AGC Ser	AAG Lys	TAT Tyr	ATA Ile	GCA Ala 200	TGG Trp	CCT Pro	TTG Leu	CAG Gln	GGC Gly 205	TGG Trp	CAA Gln	GCC Ala	624
ACG Thr 210	TTT Phe	GGT Gly	GGT Gly	GGC Gly	GAC Asp	CAT His 215	CCT Pro	CCA Pro	AAA Lys	TCG Ser	GAT Asp 220	CTG Leu	GTT Val	CCG Pro	CGT Arg	672
GGA Gly 225	TCC Ser	CCA Pro	GGA Gly	ATT Ile	CCC Pro 230	GGG Gly	TCG Ser	ACT Thr	CGA Arg	GCG Ala 235	GCC Ala	GCA Ala	TCG Ser	TGA		717

FIG.23